SOMATIC EXPANSION OF PREMUTATION ALLELES AND THE ROLE OF THE MISMATCH REPAIR AND BASE EXCISION REPAIR PROTEINS ON REPEAT EXPANSION IN A MOUSE MODEL OF THE FRAGILE X-RELATED DISORDERS

by

RACHEL ADIHE LOKANGA

Under the supervision of PROF. M. IQBAL PARKER and Dr. KAREN USDIN

Thesis presented in fulfillment of the degree of

DOCTOR OF PHILOSOPHY In MEDICAL BIOCHEMISTRY

University of Cape Town

in collaboration with

The National Institutes of Health (USA)

and

The International Centre for Genetic Engineering and Biotechnology

DECEMBER 2015

DECLARATION

"I, Rachel Adihe Lokanga, hereby declare that the work on which this thesis is based is my original work (except where acknowledgements indicate otherwise) and that neither the whole work nor any part of it has been, is being, or is to be submitted for another degree in this or any other university. I authorize the University to reproduce for the purpose of research either the whole or any portion of the contents in any manner whatsoever".

Signed by candidate

Signature Removed

Signature: Rachel Lokanga Date: 11/05/2015

DEDICATION

To my beloved parents, Okita Lokanga Yungu and Jacqueline Lokaso, who worked hard their entire lives to raise me in a loving environment,

To my daughters, Joyce and Johanne Benedict, who have encouraged me throughout the whole process of writing this thesis.

To my younger sisters, Sarah Rose Dembo and Cathy Mboyangawo, who always prayed for me and encouraged me when I was down,

To all my brothers and sisters in the faith, especially pastor Augustine Blay and brother Joseph Adih for their constant encouragement, support and for praying for me,

To my God, the Lord almighty, for giving me strength to write this thesis, for good health, and provision, I dedicate this work.

ACKNOWLEDGEMENTS

I would like to thank and express my sincere gratitude to the following individuals and organizations:

My advisor, Dr. Iqbal Parker, for accepting me in his group and for his guidance and invaluable advice throughout the execution of this project;

My co-advisor, Dr. Karen Usdin, who taught me the fundamentals of how to conduct a research project, interpret my data, and for her invaluable guidance and advice throughout this project; her support is much appreciated.

Dr. Daman Kumari, who helped me with the X inactivation assay by teaching me how set up real time PCR mixture reactions, cycle conditions and analyzing the X inactivation assay data;

Dr. Gail Seabold, for her encouragement and time to read my thesis,

Dr. Bruce Hayward for reading my thesis,

Dr. Zhao, Xiao-Nan and Mr. Zhou, Yifan, and all past lab members for their help and support throughout my entire time in the laboratory,

Dr. Huiyan Lu for help in mouse surgeries including ovariectomies, and sperm collection and all technicians at the mouse facility for taking care of the mice.

I also thank the National institutes of Health for funding my studies.

TABLE OF CONTENTS

CHAPTER 1: INTRODUCTION 1

1.1 The Fragile X-related disorders (FXDs) and the *FMR1* gene 2

1.1.1 FRAGILE X MENTAL RETARDATION GENE (*FMR1*) 4

1.1.2 *FMR1* ALLELE CLASSIFICATION 5

1.1.3 PATHOLOGY ASSOCIATED WITH *FMR1* PREMUTATION ALLELES 7

1.1.3.1 FXTAS 7

1.1.3.2 FXPOI 8

1.1.4 PATHOLOGY ASSOCIATED WITH *FMR1* FULL MUTATION ALLELES 9

1.1.4.1 *FXS* 9

1.2 Repeat expansion background 10

1.2.1 EXPANSION-PRONE REPEATS FORM SECONDARY DNA STRUCTURES 11

1.2.2 MECHANISMS IMPLICATED IN REPEAT INSTABILITY 13

1.3 MMR and BER Implication in repeat instability 14

1.3.1 THE MISMATCH REPAIR PATHWAY 14

1.3.2 BASE EXCISION REPAIR 18

1.4 The Fragile X PM Mouse model 21

1.5 Thesis Objectives/outline 22

CHAPTER 2: SOMATIC INSTABILITY OF PM ALLELES IN A HUMAN PM CELL LINE AND IN A FX PM MOUSE MODEL 24

2.1 Introduction 25

2.2 Results: 26

2.2.1 EXPANSION IS SEEN IN A HUMAN LYMPHOBLASTOID CELL LINE 26

2.2.2 THE EXTENT OF SOMATIC EXPANSION VARIES IN DIFFERENT ORGANS OF MALE MICE. 30

2.2.3 SOMATIC EXPANSION IS LESS EXTENSIVE IN FEMALE MICE 32

2.2.3.1 Factors contributing to FX PM repeat expansion 35

2.2.3.1.1 XCI contributes to the reduced SII seen in FX PM female mice 35

2.2.3.1.2 Skewed XCI does not explain differences in female SII. 38

2.2.3.1.3 Some MMR and BER proteins correlate with SII in different organs. 40

2.2.3.1.4 Lack of gender biases in the levels of DNA repair proteins in FX PM mice. 42

2.2.3.1.5 Lack of protective effect of estrogen on expansion in FX PM females mice 44

2.3 Discussion 45

CHAPTER 3: EFFECT OF A NULL MUTATION IN THE MMR GENES, *MSH2* AND *MSH6,* ON THE FX PM REPEAT EXPANSION 49

3.1 Introduction 50

3.2 Results: 52

3.2.1 SOMATIC EXPANSION OF THE FX PM ALLELES IS REDUCED IN MSH2 OR MSH6 DEFICIENT MICE. 52

3.2.1.1 Somatic expansion is eliminated in MSH2 deficient mice. 53

3.2.1.2 Somatic expansion is reduced in MSH6 deficient mice 57

3.2.2 GERM LINE EXPANSION OF FX PM ALLELES IS REDUCED IN MSH2 AND MSH6 DEFICIENT MICE. 60

3.2.2.1 Intergenerational expansion is eliminated in MSH2 deficient mice 62

3.2.2.2 Intergenerational expansion is also reduced in MSH6 deficient mice 66

3.3 Discussion 71

CHAPTER 4: EFFECT OF A MUTATION IN THE BER GENE, *POLB* ON FX PM REPEAT EXPANSION 75

4.1 Introduction 76

4.2 Results: 78

4.2.1 SOMATIC INSTABILITY IS REDUCED IN MICE WITH A Y265C MUTATION IN POLβ. 78

4.2.2 GERMLINE EXPANSION IS ALSO REDUCED IN MICE WITH A Y265C MUTATION IN POLβ 80

4.3 Discussion: 87

CHAPTER 5: CONCLUSIONS 90

5.1 Conclusions 91

CHAPTER 6: MATERIALS AND METHODS 95

6.1 Mouse maintenance and breeding 96

6.2 DNA isolation 97

6.3 Genotyping 98

6.4 Repeat analysis 101

6.5 The XCI assay 102

6.6 X inactivation ratio assay 103

6.7 Western blotting 106

6.7.1 PROTEIN EXTRACTION AND QUANTIFICATION 106

6.7.2 POLYACRYLAMIDE GEL ELECTROPHORESIS AND WESTERN BLOTTING 107

6.8 ELISA essay: 109

6.9 Statistical analysis 109

6.10 List of publications 111

6.11 References 112

ABBREVIATIONS

APE1:	apurinic/apyrimidinic endonuclease 1
ATM:	ataxia-telangiectasia, Mutated
ATP:	adenosine triphosphate
ATR:	ATM and Rad3-related
BER:	base Excision Repair
BSA:	Bovine Serum Albumin
DM1:	Myotonic Dystrophy type 1
DNA:	deoxyribonucleic acid
dRP lyase:	5'-deoxyribose-5-phosphate lyase
EDTA:	ethylenediaminetetraacetic acid
EXO1:	exonuclease 1
FAM:	6-carboxyfluorescein
FEN1:	flap endonuclease 1
FM:	Full mutation
Fmr1:	Fragile X mental retardation 1 gene
FRDA:	Friedreich ataxia
FS:	folate-sensitive
FX PM:	Fragile X premutation
FXDs:	Fragile X-related disorders
FXPOI:	Fragile X-associated primary ovarian insufficiency
FXS:	Fragile X syndrome
FXTAS:	Fragile X-associated tremor and ataxia syndrome
HD:	Huntington disease
HEX:	4, 7, 2′, 4′, 5′, 7′-hexachloro-6-carboxyfluorescein
HNPCC:	hereditary non-polyposis colorectal cancer
IDLs:	insertion-deletion
iPSCs:	induced pluripotent stem cells
KO:	Knockout
LP BER:	long patch BER
mGluR5:	metabotropic glutamate receptor 5
MLH1:	MutL protein homolog 1
MMR:	mismatch repair
MSH2:	MutS protein homolog 2
MSH3:	MutS protein homolog 3
MSH6:	MutS protein homolog 6
MutSα,	heterodimer of MSH2 and MSH6
MutSβ	heterodimer MSH2 and MSH3
ncMMR:	non-canonical MMR
NEIL1:	Nei endonuclease VIII-like

NER:	nucleotide excision repair
NHEJ:	non-homologous end joining
OGG1:	8-oxoguanine glycosylase
ORF:	open reading frame
PAGE:	polyacrylamide gel electrophoresis
PCNA:	proliferating cell nuclear antigen
PCR:	Polymerase Chain Reaction
PM:	Premutation
PolB:	DNA polymerase β (beta)
$PolB^{c}$:	Y265C mutation in *PolB*
Pol δ:	DNA polymerase δ (delta)
Polε:	DNA polymerase ε (epsilon)
RED:	Repeat Expansion Disease
RFC	replication factor C
SD	standard deviation
SDS-PAGE	sodium dodecyl sulfate polyacrylamide gel electrophoresis
SII:	somatic instability index
SBMA:	spinal and bulbar muscular atrophy
SMC1a:	Structural Maintenance Of Chromosomes 1A
SN BER:	single-nucleotide BER
SP BER:	short patch BER
SP-PCR:	small pool PCR
TCR:	Transcription Coupled Repair
UBE1	ubiquitin-like Modifier-activating enzyme 1
UTR:	untranslated region
WT:	Wild type
XCI:	X chromosome inactivation

Abstract

The Fragile X-related disorders arise from an unusual mutation in the X-linked *FMR1* gene. The mutation involves expansion, or an increase in the number of repeats, in a CGG•CCG repeat tract located in its 5' untranslated region. *FMR1* alleles carrying 55-200 repeats are called Premutation (PM) alleles, and cause Fragile X associated tremor/ataxia syndrome (FXTAS) and Fragile X-associated primary ovarian insufficiency (FXPOI). *FMR1* alleles having more than 200 repeats are referred to as full mutation (FM) alleles and cause Fragile X syndrome (FXS). These different alleles arise by intergenerational expansion of the repeat tract from smaller unstable alleles by a mechanism that is unknown.

We have shown that in addition to germ line expansion, somatic expansion also occurs in a human cell line *in vivo* and in a FX PM mouse model. In the mouse model, we found that the extent of somatic instability is dependent on age, gender and tissue. Specifically, organs such as brain, liver and gonads are susceptible to expand more than heart and kidney and expansion is much more frequent in males than in females.

No differences were found between male and female mice in the levels of the DNA repair proteins that had already been implicated in repeat expansion in model systems of other disorders thought to arise via a similar mechanism. Neither were there any differences between males and females in the amounts of proteins produced from X-linked DNA repair genes. We also showed that estrogen did not protect against expansion. However, we found that PM alleles expanded exclusively when they were located on the active X chromosome. Thus some of the differences between males and

females in the level of somatic expansion might be due to the fact that females undergo X inactivation and thus have the PM allele on the inactive X chromosome in half (~50%) of their cells. It also indicates that transcription and/or an open chromatin configuration is required for expansion in the FX PM mouse.

We also examined the effect of mutations in the DNA repair genes *Msh2*, *Msh6* and *PolB*, on intergenerational and somatic expansions in the FX PM mouse model. The *Msh2* gene product, MSH2, interacts with MSH6 to form MutSα, one of the two MSH2-containing complexes found in mammalian cells. All expansions were abolished in *Msh2* null mice and in *Msh6* null mice expansions were significantly decreased. This suggests a role for MutSα in the repeat expansion process. Since we found that the *PolB* mutation we tested was embryonic lethal in the homozygous state, we examined the extent of paternal germ line and somatic expansion in heterozygous males. We found that even in these animals there was a significant reduction in both germ line and somatic expansions. Since *PolB* is involved in the Base Excision Repair pathway and MutSα is involved in the Mismatch Repair pathway, our data suggest a model in which these pathways interact to generate expansions.

Chapter 1: Introduction

1.1 The Fragile X-related disorders (FXDs) and the *FMR1* gene

The Fragile X-related disorders (FU *et al.* 1991; VERKERK *et al.* 1991; YU *et al.* 1992; ALLINGHAM-HAWKINS *et al.* 1999; HAGERMAN *et al.* 2001; HAGERMAN *et al.* 2004) belong to a larger category of disorders known as the Repeat Expansion Disorders. These disorders comprise approximately 40 human inherited neurological, neuromuscular or neurodegenerative ailments that arise from the expansion of an unstable microsatellite repeat tract within a specific gene (PEARSON *et al.* 2005). These microsatellite repeat tracts can be of variable sequence and size and can be found in different locations within the affected gene, including the 3' or 5' untranslated regions (UTRs), introns and exons (Figure 1.1). These disorders have been known for decades; but their genetic basis was first described in 1991 when the expansion of a specific trinucleotide repeat was revealed to be the most common cause of Spinal and bulbar muscular atrophy (SBMA), a progressive neuromuscular disorder that results in muscle atrophy, and of Fragile X syndrome, one of the Fragile X related disorders and the leading cause of inherited mental disability (LA SPADA *et al.* 1991; VERKERK *et al.* 1991). Many other repeat expansion disorders have since been identified and the list of disease-causing repeats expanded to include not just trinucleotide repeats which constitute the majority of these disorders, but also tetra, penta, and even dodecameric nucleotide repeats (LALIOTI *et al.* 1997; MATSUURA *et al.* 2000; LIQUORI *et al.* 2001; PEARSON *et al.* 2005; MIRKIN 2006). For each of these disorders, expansions are confined to a single repeat tract in a disease-specific gene.

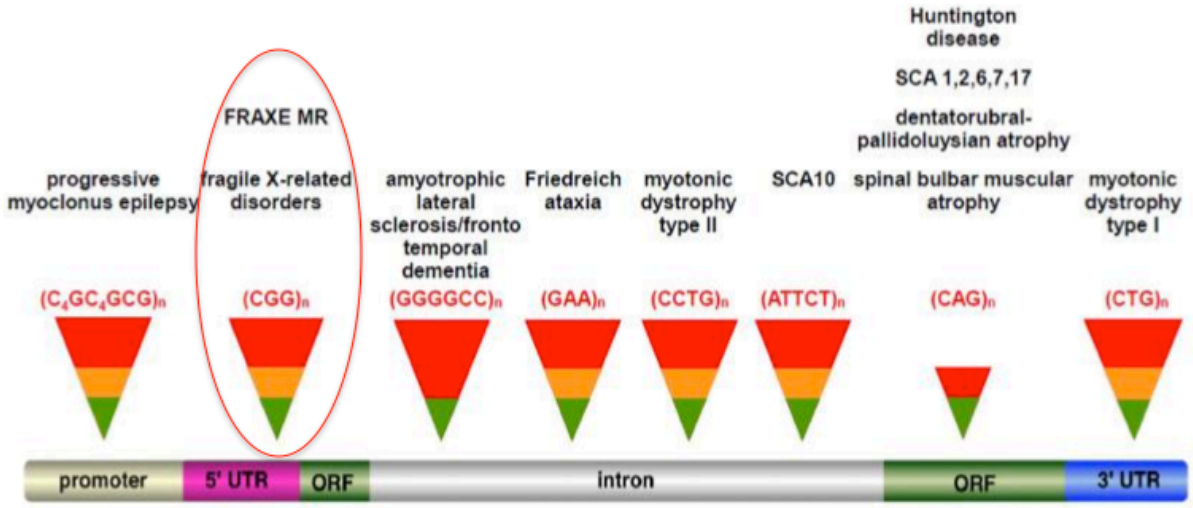


Figure 1.1 Repeat Expansion Diseases in humans. Illustration of different nucleotide repeats that cause diseases and their locations on a generic gene. Green represents normal alleles, yellow signifies carriers of premutation length and red denotes disease length alleles (KUMARI *et al.* 2012).

The Fragile X related disorders (FXDs) include FX-associated primary ovarian insufficiency (FXPOI), Fragile X-associated tremor/ataxia syndrome (FXTAS), and Fragile X syndrome (FXS) (FU *et al.* 1991; VERKERK *et al.* 1991; YU *et al.* 1992; ALLINGHAM-HAWKINS *et al.* 1999; HAGERMAN *et al.* 2001; HAGERMAN *et al.* 2004). They all arise from the expansion of an unstable CGG•CCG-repeat tract in the 5′ UTR of the *Fragile X mental retardation 1* gene (*FMR1*). They are so called because of a fragile site (FS), an apparent constriction or break, in the chromosome that colocalizes with the *FMR1* gene when cells are propagated under folate stress (HECHT AND SUTHERLAND 1985; VERKERK *et al.* 1991)(Figure 1.2).

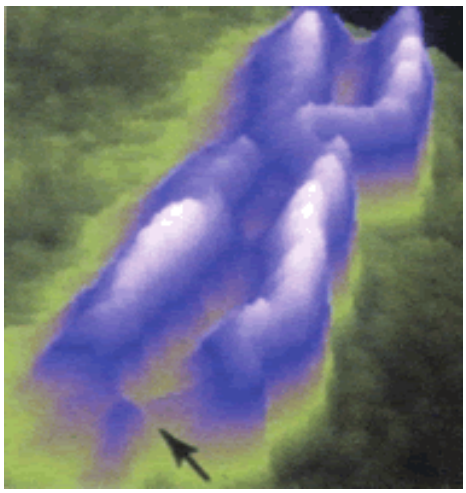


Figure 1.2 Location of a fragile site on X chromosome. Atomic force microscopic image of a Fragile X chromosome showing a constriction (FS) on the X chromosome. The Arrow indicates the fragile site location.
Photo Credit: Dr. Ben Oostra, Wellcome Images

1.1.1 Fragile X mental retardation gene (*FMR1*)

FMR1 is a gene that is found on the long (q) arm of the X chromosome at 27.3 (VERKERK *et al.* 1991). As an X-linked gene, it is subject to X chromosome inactivation (XCI), a phenomenon that results in random silencing of one of the two copies of the X chromosomes (maternal or paternal) found in females (LYON 1961; KIRCHGESSNER *et al.* 1995; AVNER AND HEARD 2001). This inactivation ensures that X-linked genes are expressed at similar level in both males and females, a phenomenon known as dosage compensation. The *FMR1* gene comprises 17 coding exons and can be alternatively spliced, resulting in different mRNAs and protein isoforms (ASHLEY *et al.* 1993; EICHLER *et al.* 1993; VERHEIJ *et al.* 1993; VERKERK *et al.* 1993; HUANG *et al.* 1996; FU *et al.* 2015).

The *FMR1* gene product is called Fragile X mental retardation protein 1 (FMRP), an RNA binding protein (ASHLEY *et al.* 1993; SIOMI *et al.* 1993) that is expressed in the majority of mammalian tissues including human and mouse tissues but has elevated expression in neurons and gonads (GIBSON *et al.* 1993; VERHEIJ *et al.* 1993; BROWN *et al.* 1998; BAKKER *et al.* 2000; OOSTRA AND WILLEMSEN 2009). In mammalian cells, FMRP is primarily cytoplasmic although it can also occasionally be present in the nucleus (DEVYS *et al.* 1993; WILLEMSEN *et al.* 1996). FMRP is believed to negatively regulate the translation of dendritic mRNAs thus offsetting the effect of mGluR5 activation in the post-synaptic neuron (COMERY *et al.* 1997; BEAR 2005).

1.1.2 *FMR1* allele classification

The length of the CGG•CCG repeat tract in the 5' UTR of the *FMR1* gene varies in the unaffected population but the most common allele has 30 repeats. Alleles with this repeat number have a negligible risk of expansion and are stable on intergenerational transmission. However, CGG•CCG repeats become unstable with increasing repeat number. *FMR1* alleles with ~ 41-54 repeats are categorized as "grey zone" (GZ) or intermediate alleles (ZHONG *et al.* 1996; NOLIN *et al.* 2003; TERRACCIANO *et al.* 2004). Unlike normal alleles, intermediate alleles are slightly prone to expand on intergenerational transmission. They were classified as intermediate alleles based on their lower risk of expansion to cause FXS in one single generation (TERRACCIANO *et al.* 2004). *FMR1* alleles having between 55 to 200 CGG•CCG repeats are named premutation (PM) alleles and confer a risk of FXTAS and FXPOI. PM alleles are unmethylated and are hyperexpressed (TASSONE *et al.* 2007b). However, the expression

of FMRP is moderately reduced in the PM alleles despite the fact that transcription is elevated. It has been suggested that the decreased levels of FMRP are likely due to inefficient translation of the transcript when it contains large repeat numbers (FENG *et al.* 1995; PRIMERANO *et al.* 2002; CHEN *et al.* 2003).

FMR1 alleles carrying over 200 CGG•CCG repeats are classified as full mutation (FM) alleles. In contrast to PM alleles, FM alleles are transcriptionally silenced. While many of these alleles still make some *FMR1* mRNA, little or no FMRP is produced (PIERETTI *et al.* 1991; FENG *et al.* 1995). The absence of FMRP results in FXS (Figure 1.3) (VERKERK *et al.* 1991).

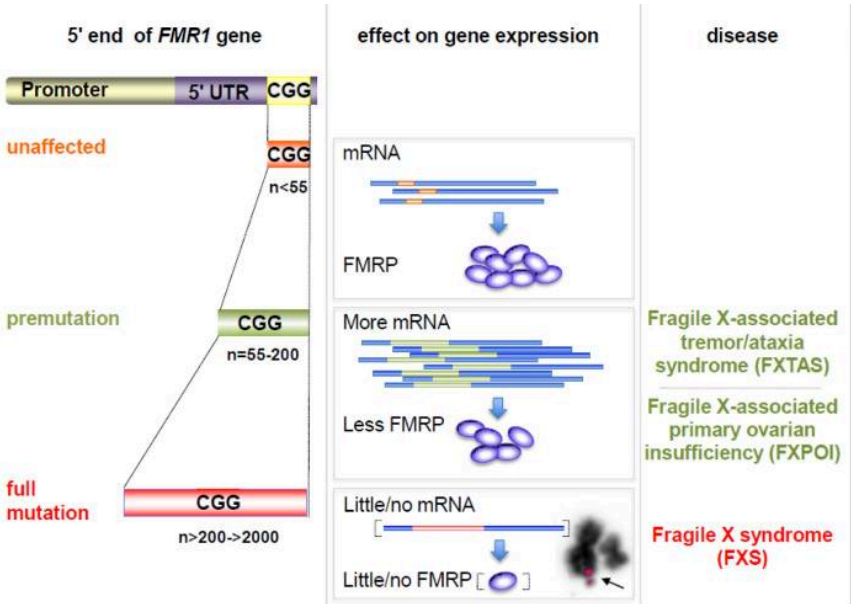


Figure 1.3 Relationship between repeat number, expression of *FMR1* mRNA, FMRP, and associated diseases. PM alleles have 55 to 200 repeats and express elevated *FMR1* mRNA levels. The level of FMRP is moderately decreased. These alleles confer an increased risk of developing FXTAS and FXPOI. Full mutation alleles have > 200 repeats. FMR1 transcription is severely reduced, resulting in no FMRP expression. Loss of FMR1 product, FMRP causes fragile X syndrome. Adapted from (KUMARI *et al.* 2012)

1.1.3 Pathology associated with *FMR1* premutation alleles

1.1.3.1 FXTAS

FXTAS is a neurological illness associated with progressive intention tremor, Parkinsonism, high blood pressure, cerebellar gait ataxia, sleep apnea, dementia, working memory impairment, and executive dysfunction and other signs including peripheral neuropathy and autonomic dysfunction (HAGERMAN *et al.* 2001; JACQUEMONT *et al.* 2003; BREGA *et al.* 2008; HAMLIN *et al.* 2011; HAMLIN *et al.* 2012; HAGERMAN 2013; HAGERMAN AND HAGERMAN 2013; USDIN *et al.* 2014). Patients with FXTAS show characteristic radiological findings specifically, increased bilateral signal intensities of the middle cerebellar peduncles on T2-weighted Magnetic resonance imaging (MRI) as well as other abnormalities such as enlarged ventricles, brain atrophy and white matter changes (BRUNBERG *et al.* 2002; JACQUEMONT *et al.* 2003; COHEN *et al.* 2006; ADAMS *et al.* 2007). In addition, characteristic intranuclear inclusions were reported in the brain and testes of affected males (GRECO *et al.* 2002; GRECO *et al.* 2007). These inclusions were shown to contain *FMR1* mRNA and numerous proteins (TASSONE *et al.* 2004; IWAHASHI *et al.* 2006). Both males and females are affected by FXTAS. However, symptoms of FXTAS are less severe in females than males due at least in part to XCI (BERRY-KRAVIS *et al.* 2005). The mechanism responsible for FXTAS pathology is not well understood, but is thought to arise from some detrimental consequence of the hyper expression FMR1 mRNA that contains a large CGG-repeat tract (HAGERMAN *et al.* 2004; HAGERMAN 2013). Two current models for disease pathology are favored, one in which the repeats sequester CGG•CCG binding proteins and one in which the repeats

trigger translation at a non-canonical translation initiation codon, a process known as repeat associated non-AUG (RAN) translation, thereby producing a toxic protein (JIN *et al.* 2007; SELLIER *et al.* 2010; RENOUX AND TODD 2012; SELLIER *et al.* 2013)

1.1.3.2 FXPOI

FXPOI is a form of ovarian pathology that affects about 20% of female carriers of PM alleles (ALLINGHAM-HAWKINS *et al.* 1999; SHERMAN 2000; SULLIVAN *et al.* 2011; HAGERMAN AND HAGERMAN 2013). This malady is characterized by fertility problems, menstrual cycle irregularities and an earlier than normal menopause. A non-linear relationship was found between the numbers of CGG•CCG repeats and the age of menopause (ENNIS *et al.* 2006). PM carriers having ~80-100 repeats have the highest risk of developing premature ovarian failure, and thus early menopause. In addition, It was found that even those PM carriers who did not have FXPOI were likely to undergo menopause 5 years earlier than the normal population with normal *FMR1* alleles (SHERMAN 2000). Therefore, FXPOI is a significant public health concern as premature menopause as been shown to increase the risk of mortality at least five fold for neurological disorders (RIVERA *et al.* 2009; SCOTT *et al.* 2014), and increase the risk of a variety of other detrimental conditions including ischemic stroke (ROCCA *et al.* 2012). As with FXTAS, FXPOI is thought to be the result of some deleterious expression of the mRNA containing a long CGG-repeat tract.

1.1.4 Pathology associated with *FMR1* full mutation alleles

1.1.4.1 *FXS*

FXS is the leading common heritable cause of intellectual disability. Its prevalence in the normal population is approximately 1 in 4000 males and 1 in 5000-8000 females (LYONS *et al.* 2015). It was first documented in 1943 by Martin and Bell as a X-linked mental retardation disorder (MARTIN AND BELL 1943). Symptoms of FXS include a learning disability that is often severe, developmental delay, hyperactivity, depression and anxiety (CORNISH *et al.* 2001). FXS is associated with autistic features and in fact is the major known monogenic source of autism (HATTON *et al.* 2006). In addition, patients affected with FXS often exhibit distinctive physical features such as a prominent jaw, large ears, large testes and elongated face (BUTLER *et al.* 1993). While both men and women are affected with FXS, females tend to be less affected because they are heterozygous and that the second X chromosome confers some protective effect on them (DE VRIES *et al.* 1996; BENNETTO *et al.* 2001; BERRY-KRAVIS *et al.* 2005). The full mutation usually leads to the silencing of the *FMR1* gene and the absence of FMRP (PIERETTI *et al.* 1991; DEVYS *et al.* 1993; VERHEIJ *et al.* 1993). Some non-repeat associated *FMR1* mutations, including missense and splice site mutations as well as deletions also cause FXS (DE BOULLE *et al.* 1993; HAMMOND *et al.* 1997; PIMENTEL 1999).

The mechanism by which CGG•CCG repeats expand from a normal allele to a PM allele or from a PM to a FM allele is unknown. However, numerous risk factors for expansion have been identified in humans as well as in mouse models. (SHERMAN *et al.*

1984; OBERLE *et al.* 1991; ROUSSEAU *et al.* 1991). The size of repeat is one of the most important risk factors for expansion. The larger the PM allele the more likely it is to expand to a FM in one generation (HEITZ *et al.* 1992; YU *et al.* 1992; FISCH *et al.* 1995). Parental gender is also an important factor. While smaller alleles are more likely to expand on paternal transmission, transmission of the FM allele from a PM allele occurs exclusively on maternal transmission (SHERMAN *et al.* 1984; OBERLE *et al.* 1991; ROUSSEAU *et al.* 1991). This is likely due to contraction of large expansions in sperm (MALTER *et al.* 1997). Beyond the repeat size and the gender of the transmitting parent, population and family studies have reported other factors that play a role in FX repeat instability including the presence, number and location of AGG interruptions as well as potential cis and trans-acting factors related to haplotype background (SNOW *et al.* 1994; NOLIN *et al.* 2003; NOLIN *et al.* 2013; LATHAM *et al.* 2014).

1.2 Repeat expansion background

Instability in Repeat Expansion Diseases shows a strong expansion bias, although contractions of the repeat are also observed (DE BOULLE *et al.* 1993; DE GRAAFF *et al.* 1995; ENTEZAM *et al.* 2007). The tendency of the repeat to expand on intergenerational transmission results in the phenomenon of anticipation in which the disease severity increases with successive generations and an earlier age of onset of symptoms is seen.

In addition to repeat expansion on intergenerational transmission, expansion has also been reported to occur in somatic cells throughout the life span of an individual in many Repeat Expansion Diseases. Several of these diseases show increased levels of somatic instability particularly for bigger alleles (TELENIUS *et al.* 1994; MORALES *et al.* 2012). The expansion of the unstable repeat in somatic cells is generally age-dependent. However, the pattern of repeat expansion is not uniform in different organs of the same individual with some organs being more prone to expansion than others (TELENIUS *et al.* 1994). Recent work in some Repeat Expansion Diseases suggests that somatic expansion can contribute to the age of the disease onset and severity of the symptoms (SWAMI *et al.* 2009). In the Fragile X-related disorders (FXDs) somatic mosaicism is often seen (PIERETTI *et al.* 1991; PRETTO *et al.* 2014). However, whether this results from somatic expansion of smaller alleles or contraction of larger ones is unknown.

1.2.1 Expansion-prone repeats form secondary DNA structures

The disease-associated repeats are able to adopt unusual conformations, including triplexes, tetraplexes, slipped strand structures and hairpins (Figure 1.4). These structures are believed to form whenever the DNA becomes single-stranded such as during transcription or replication (MIRKIN 2006). Single-stranded DNA containing CNG repeats, where “N” is any nucleotide, can fold into hairpin-like structures made up of a combination of Watson-Crick base pairs and mismatched base pairs in a 2:1 ratio (Figure 1.4a) (GACY *et al.* 1995; MITAS *et al.* 1995; USDIN AND WOODFORD 1995; USDIN 1998)

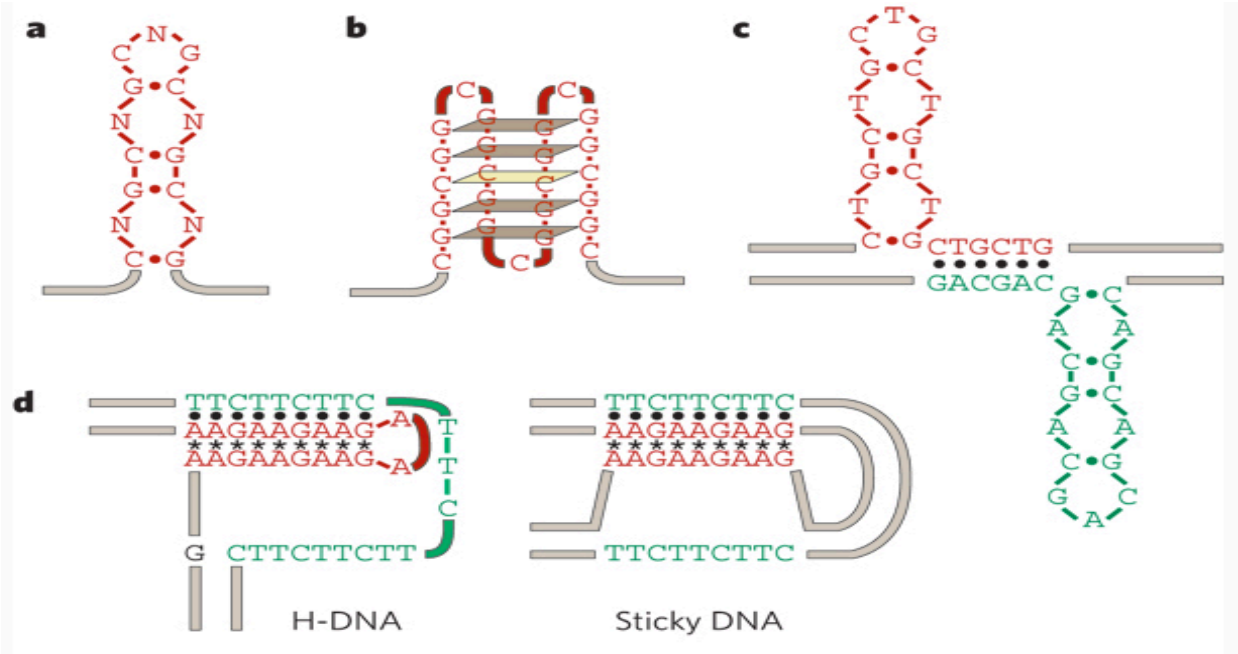


Figure 1.4 Different DNA conformations formed by unstable repeats. (a) Hairpin structure made-up of (CNG)n repeat. (b) G-quartets conformation formed by (CGG)n repeats (c) Slipped (S) DNA (d) Different triplexes made with (GAA)n repeats (other conformations of sticky DNA are not shown) Adapted from (MIRKIN 2006)

.

In addition to hairpin-like structures, the individual strands of the GC-rich repeat, including the CGG and CCG can also adopt other unusual DNA conformations (USDIN AND WOODFORD 1995). For example, a quadruplex-like structure (G-quartets structure) (Figure 1.4b) can be formed when an unpaired DNA strand of CGG•CCG repeats fold into a tetra-helical structure that is stabilized by G quartet motifs (MITAS *et al.* 1995; USDIN AND WOODFORD 1995; PEARSON AND SINDEN 1996; MIRKIN 2007). Slipped DNA structures can arise when the two DNA strands reanneal out of register resulting in hairpins or loopouts on both strands (Figure 1.4c). GAA•TTC repeats can also form different sorts of intramolecular triplexes (1.4d), a type of three-stranded DNA structure (GACY *et al.* 1998).

Interruptions to the purity of the repeat tract reduce the stability and length of the secondary structures that can be formed. These interruptions decrease the risk of expansion in many Repeat Expansion Diseases (EICHLER *et al.* 1995). For example, in the FXDs, AGG interruptions increase the stability of the repeat tract during maternal transmission (NOLIN *et al.* 2003).

1.2.2 Mechanisms implicated in repeat instability

The mechanism(s) responsible for repeat expansion and contraction on either intergenerational transmission or in somatic tissues are poorly understood. However, several hypotheses have been proposed for both expansion and contraction. Most of these hypotheses are founded on the idea that unusual secondary DNA structures formed by the expansion-prone repeats result in instability as a result of problems that arise during different DNA metabolic processes including DNA replication, recombination or repair. However, evidence from several model systems and affected humans suggests that expansion does not result from a problem with chromosomal DNA replication per se. For example, in diseases like Myotonic Dystrophy type 1 (DM1), Huntington disease (HD), and Friedreich ataxia (FRDA), expansions are seen in organs with a low proliferative capacity such as kidney, as well as in non-dividing cells such as neurons or gametes (TELENIUS *et al.* 1994; THORNTON *et al.* 1994; LIA *et al.* 1998; DE BIASE *et al.* 2007; VAN DEN BROEK *et al.* 2007; GONITEL *et al.* 2008). In addition, in a mouse model for HD, expansions are seen in post-meiotic, haploid germ cells (KOVTUN AND MCMURRAY 2001). In addition, in both the FXDs and DM1, there is a strong maternal age effect on expansion risk (YRIGOLLEN *et al.* 2012; MORALES *et al.* 2015).

This indicates that expansions in these disorders can occur in the oocyte, another cell type that does not divide. These data favor the hypothesis that expansion arises from certain problems during DNA repair or recombination rather than chromosomal DNA replication per se.

Many different DNA repair pathways operate in mammalian cells. These pathways include base excision repair (BER), homologous recombination (HR), mismatch repair (MMR), nucleotide excision repair (NER), non-homologous end joining (NHEJ) (SANCAR *et al.* 2004). While work in vitro, and in bacteria and yeast, suggest that many of these processes can in theory produce expansions; MMR and BER are the major pathways that have been suggested to cause expansion in mouse models.

1.3 MMR and BER Implication in repeat instability

1.3.1 The mismatch repair pathway

MMR is the major pathway for the repair of base substitution mismatches and insertion-deletion mismatches (IDLs) introduced during DNA replication in prokaryote and eukaryote organisms (KOLODNER AND MARSISCHKY 1999; IYER *et al.* 2006). Some of the MMR components participate in a number of other processes of DNA repair including homologous recombination, somatic hypermutation and interstrand-crosslink repair.

In mammalian cells, MSH2, a key component of the MMR pathway, hetero-dimerizes with either MSH6 to form a MutSα complex or with MSH3 to form a MutSβ complex. MMR repair is initiated by recognition of the mismatch by these

heterodimers (KUNKEL AND ERIE 2005). The recognition properties of MutSα are different from that of MutSβ heterodimers, with MutSα preferentially recognizing base-base mismatches and insertion/deletion mismatches (IDL) of 1-2 nucleotides, while MutSβ can also recognize insertion/deletions (IDL) involving several extra nucleotides (MCCULLOCH *et al.* 2003).

After recognition of the mismatch, a MutL complex is then recruited (LI 2008). Mammalian cells contain 3 MutL complexes: MutLα , MutLβ, and MutLγ . These complexes are heterodimers involving MLH1 and PMS1, PMS2 or MLH3. While MutLα (MLH1-PMS2) is known to play an important role in repairing various mismatches, MutLγ (MLH1-MLH3) is involved in repairing IDLs (FLORES-ROZAS AND KOLODNER 1998) and involved in meiotic recombination (SANTUCCI-DARMANIN AND PAQUIS-FLUCKLINGER 2003). The role of MutLβ (MLH1-PMS1) is not yet understood (KOLODNER AND MARSISCHKY 1999). The MutS/MutL complex then interacts with the proliferating cell nuclear antigen (PCNA) (CLARK *et al.* 2000). This interaction leads to the activation of the endonuclease activity of MutL that results in the formation of a nick in the vicinity of the mismatch. Exonuclease 1 (EXO 1) is then recruited to the nick and creates a single-stranded gap. This gap is then filled by DNA polymerase δ (Polδ), and the MMR is completed when the nick is ligated using DNA ligase I (Lig1) (KUNKEL AND ERIE 2005; LI 2008).

Mutations in MMR genes increase genome wide microsatellite instability (MSI) and are the leading cause of many hereditary and sporadic cancers in humans (MODRICH AND LAHUE 1996; KOLODNER AND MARSISCHKY 1999; POYNTER *et al.* 2008). For

example, germ line mutations of the MMR genes *MSH2*, *MSH6*, *MLH1* and *PMS2* have been linked to hereditary cancers including hereditary non-polyposis colorectal cancer (HNPCC) or Lynch syndrome (LYNCH *et al.* 1985; ABDEL-RAHMAN *et al.* 2006; RAHNER *et al.* 2013). While mutations in key MMR genes result in increased MSI, work in various model systems of some Repeat Expansion Diseases has shown variable effects of mutations in different genes involved in MMR on repeat expansion.

For instance, in a mouse model of DM1, loss of MSH2 eliminated repeat expansion in both germline and somatic cells (VAN DEN BROEK *et al.* 2002; SAVOURET *et al.* 2003), while in a transgenic mouse model of HD loss of MSH2 eliminated somatic expansions and paternal but not maternal expansions (MANLEY *et al.* 1999b; KOVTUN AND MCMURRAY 2001; WHEELER *et al.* 2003). However, in a mouse model of Friedreich ataxia (FRDA), MSH2 was shown to protect against intergenerational repeat contractions (EZZATIZADEH *et al.* 2012), but contributed to somatic expansion in these mice (BOURN *et al.* 2012). In induced pluripotent stem cells (iPSCs) from FRDA patient fibroblasts cells, shRNA silencing of *MSH2* prevented repeat expansion (KU *et al.* 2010; DU *et al.* 2012).

The effect of the loss of MSH3 and MSH6 on repeat expansion was also examined in various mouse models of some Repeat Expansion Diseases. While the loss of MSH3 significantly decreased all repeat expansions in the mouse models of DM1 (FOIRY *et al.* 2006), in a HD model loss of MSH3 decreased somatic expansions but not germ line expansions (DRAGILEVA *et al.* 2009), while in a FRDA mouse model MSH3 was shown to protect against contractions. Loss of MSH6 did not reduce either

germ line or somatic expansions in either HD or DM1 mouse models (VAN DEN BROEK *et al.* 2002; FOIRY *et al.* 2006; DRAGILEVA *et al.* 2009). In contrast, in a mouse model of FRDA, MSH6 was shown to protect against both expansions and contractions in the germline but contributed to expansions in somatic cells (BOURN *et al.* 2012; DU *et al.* 2012).

The molecular basis by which the MMR proteins act to promote repeat expansion remains to be elucidated. One suggestion is that MutSβ promotes expansion by its ability to bind and stabilize the hairpin structure instead of processing the mismatches by the MMR pathway as illustrated in figure 1.5 (OWEN *et al.* 2005). This model, which is also referred to as the "hijacking" model is based on the fact that binding of MutSβ to secondary structures formed by repeats inhibits its ATPase activity. It has been proposed that this leads to defects in downstream signaling that in turn result in error-prone DNA repair that result in incorporation of unrepaired loops as expansions (OWEN *et al.* 2005; MCMURRAY 2008; LANG *et al.* 2011). However, this model was challenged when subsequent studies with slightly different substrates did not show such inhibition (TIAN *et al.* 2009). Furthermore, it was also shown that the ATPase domain is required for expansion in a DM1 mouse model (TOME *et al.* 2009). This finding suggests that the role of MMR in promoting expansion is not confined to the recognition or binding of hairpins by MutSβ (TOME *et al.* 2009). This idea is strengthened by recent work that implicates MutLα and MutLγ in the expansion process (GOMES-PEREIRA *et al.* 2004; PINTO *et al.* 2013; EZZATIZADEH *et al.* 2014).

1.3.2 Base Excision Repair

Base Excision Repair (BER) was discovered about ~40 years ago by Tomas Lindahl (LINDAHL 1974; FRIEDBERG AND LINDAHL 2004). This pathway is responsible for repairing base lesions that are caused by oxidation, deamination, alkylation, depurination or depyrimidination (ROBERTSON *et al.* 2009). The BER pathway is initiated by the recognition of a damaged base by an appropriate DNA glycosylase. This DNA glycosylase catalyzes the cleavage of an N-glycosidic bond, effectively excising the damaged base that leads to the creation of an apurinic or apyrimidinic site (AP site) (LU *et al.* 2001). The AP site is subsequently processed by the AP Endonuclease 1(APE1), an endonuclease that cleaves the DNA phosphodiester backbone and generates a single-stranded DNA nick 5' to the AP site. The AP site can also be processed by a DNA AP lyase of some DNA glycosylases creating a nick 3' to the AP site.

The processing of the AP site leads to creation of single nucleotide gap in the DNA that contains a 3' hydroxyl and 5' deoxyribose phosphate group (dRP) (ASPINWALL *et al.* 1997; DIANOV *et al.* 2003). The gap is processed via the single nucleotide BER (SN-BER) pathway when the 5' deoxyribose phosphate group is not oxidized. In this case, DNA polymerase beta (Polβ), a key enzyme of the BER pathway that possesses both 5′dRP lyase and gap-filling activities, inserts a single nucleotide with the resultant nick being sealed by DNA ligase I (Lig 1) (LIU AND WILSON 2012). However, if the AP site or the 5' phosphate group is oxidized or reduced, an alternate sub-pathway, long patch BER (LP-BER), is used for repair.

In vitro work has led to the suggestion that LP-BER could lead to expansion via a "hit and run" mechanism (LIU *et al.* 2005) and as illustrated in figure 1.5B(USDIN *et al.* 2015),involving the coordination of Polβ and Flap Structure-Specific Endonuclease 1 (FEN1) in a process of stepwise single-nucleotide gap-filling synthesis by Polβ with the resultant small flaps being partially removed by FEN1 to generate a ligatable 5' end (LIU *et al.* 2005). FEN1 is a DNA repair protein that plays an essential role in maintaining the genome stability by removing DNA flaps during DNA replication and repair (KLUNGLAND AND LINDAHL 1997). An alternate pathway involving the coordinated processing of slipped hairpins by a combination of Polβ and Polδ has also been suggested as a source of expansions (CHAN *et al.* 2013).

Evidence that BER generates expansions has come from the demonstration that OGG1 (KOVTUN *et al.* 2007) and NEIL1 (MOLLERSEN *et al.* 2012), DNA glycosylases involved in the early steps of BER reduce the extent of somatic expansion. However, loss of OGG1 did not affect germ line expansions while the loss of NEIL1 reduced the average size of the expansions but not their frequency. This has led to the suggestion that the contribution of BER to the expansion process may be limited to the generation of nicks that are then "hijacked" by the MMR proteins to generate the expansions in some way.

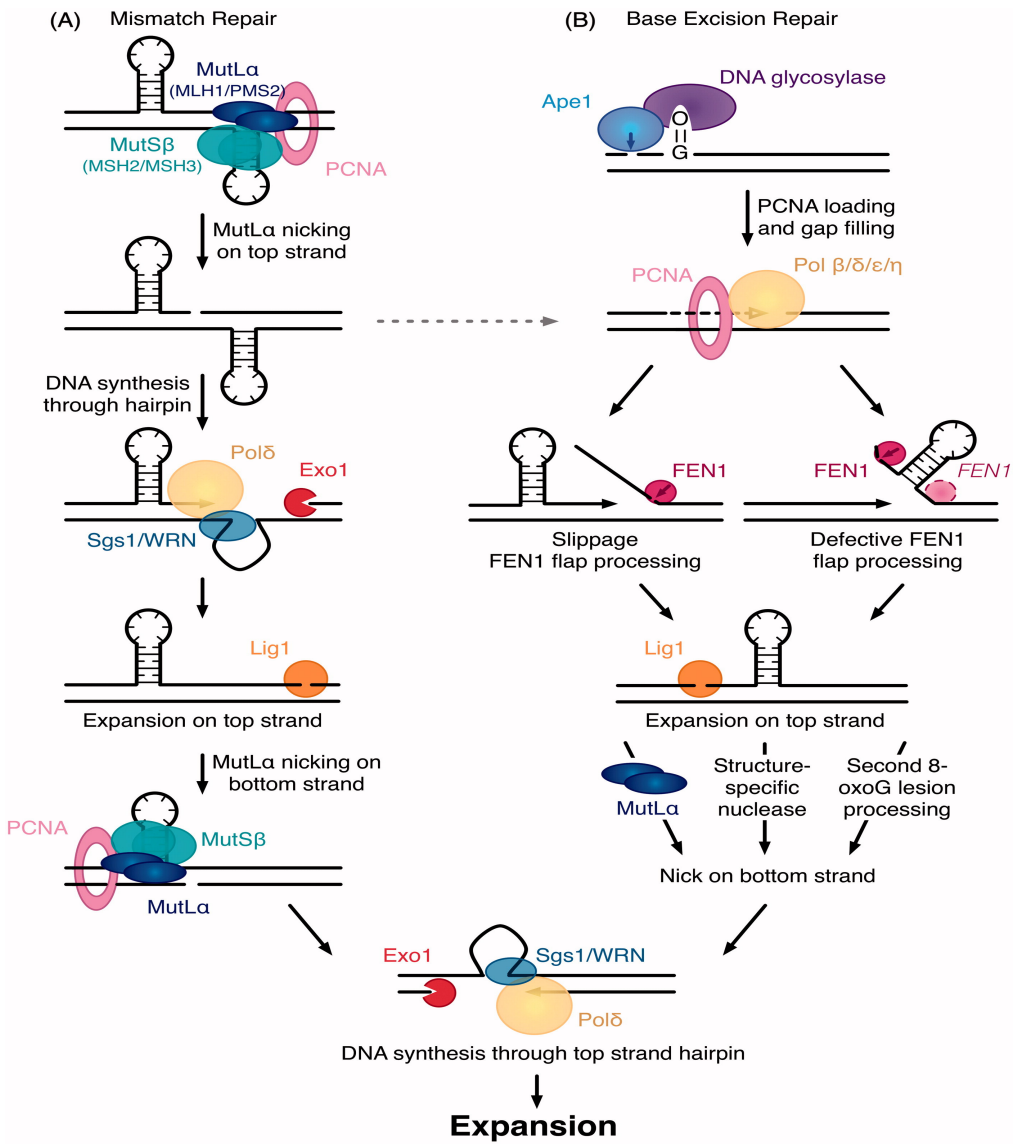


Figure 1.5 Expansion models involving both MMR and BER. (A) MMR pathway: MutS complexes bind to the hairpin formed by the repeats. PCNA and MutLα also bind to MutS complexes to form a ternary complex. MutLα nicks the top or bottom strand. This is followed by the generation of a gap by EXO1 that extends past the hairpin, DNA synthesis to fill the gap and then sealing of the nicks by Lig1. The repair leads to an expansion on both strands. (B) BER pathway: DNA glycosylase binds to oxidized base, followed by creation of a nick by Ape1. Polβ alone or in combination with Polδ, Polε, or Polη performs gap filling DNA synthesis that can result in expansion (left pathway). Hairpins can obstruct FEN1 cleavage (shown by pink dotted FEN1), preventing flap cleavage (shown by red solid FEN1) and leading to expansions as well. Expansions occur on both strands (top and bottom). Adapted from (USDIN *et al.* 2015).

1.4 The Fragile X PM Mouse model

Several transgenic mouse models have been generated to study the mechanism of CGG•CCG repeat instability. However, these mouse models showed little repeat instability upon intergenerational transmissions (LAVEDAN *et al.* 1997; BONTEKOE *et al.* 2001). To attempt to model the mechanism of repeat instability in mice, the Usdin laboratory generated a PM mouse model in which the endogenous murine repeat tract was replaced by a fragment containing ~130 CGG•CCG repeats (ENTEZAM *et al.* 2007). This mouse model recapitulates several key features seen in human carriers of PM alleles including the elevated levels of *Fmr1* mRNA, the reduced levels of FMRP, the presence of ubiquitin-positive intranuclear inclusions in the brain, Purkinje cell loss, as well as a high frequency of intergenerational expansion (ENTEZAM *et al.* 2007).

Using this mouse model, the Usdin laboratory has shown that mutations in *ATM* (Ataxia-Telangiectasia, Mutated) and *ATR* (ATM and Rad3-related), the two kinases responsible for regulating the response to DNA damage and stalled replication forks increase repeat expansion (ENTEZAM AND USDIN 2008; ENTEZAM AND USDIN 2009). Furthermore, they have showed that oxidative damage exacerbates germ line repeat expansion (ENTEZAM *et al.* 2010). They also showed that Transcription Coupled Repair (TCR) is not involved in generating repeat expansion in the FX PM mice (ZHAO AND USDIN 2014).

1.5 Thesis Objectives/outline

The study described in this thesis attempts to shed light on the expansion mechanism responsible for the Fragile X-related disorders. The work presented extends previous work in the Usdin laboratory aimed at understanding the somatic instability of the PM alleles along with genetic factors that impact both germ line and somatic repeat expansion(ENTEZAM AND USDIN 2008; ENTEZAM AND USDIN 2009).

Transcription or an open chromatin conformation facilitates expansion in FX PM mice.

A previous doctoral student in the Usdin laboratory observed somatic expansion of the CGG•CCG repeat in FX mice that seemed more extensive in males than in females. However, it was not clear why some organs were more prone to expand than others and why somatic instability was less prevalent in females. Chapter 2 describes my examination of CGG•CCG repeat instability in different organs of male and female mice and of different potential expansion risk factors that might explain the tissue differences and effects of gender. The results presented indicate that a major contributor to the gender bias in the expansion frequency was the fact that expansion does not occur on the inactive X chromosome. This suggests that expansion of the CGG•CCG repeat in PM mice requires transcription or an open chromatin configuration. However, XCI does not explain all the gender difference and we did not identify any gender-specific differences in expression of DNA repair proteins that could account for the residual excess of expansions in males. However, tissue differences in the levels of

various key DNA repair proteins did suggest a basis for some of the differences in the propensity of different organs to show expansions.

The Mismatch Repair proteins MSH2 and MSH6 promote expansion in FX PM mice

The role that MMR proteins play in CGG•CCG repeat expansion was not known. Evidence presented in chapter 3 and 4 suggests that MutSα protects the genome against repeat contractions. It also promotes repeat expansions, but does so in a MutSβ-dependent way.

***PolB* also facilitates expansion of the PM alleles**

Previous work in the Usdin laboratory indicated that potassium bromate, a powerful oxidizing agent, exacerbated germ line expansion in the CGG•CCG repeat, suggesting the role of oxidative damage in promoting expansion. Results presented in chapter 5 demonstrate that central events in the BER pathway, the major pathway by which oxidative damage is repaired in mammalian cells, is also involved in CGG•CCG repeat expansion (MOLLERSEN *et al.* 2012).

Chapter 2: Somatic instability of PM alleles in a human PM cell line and in a FX PM mouse model

2.1 Introduction

A hallmark of all Repeat Expansion Disorders is the propensity of the disease-associated repeat to undergo expansion in the germ line (HEITZ *et al.* 1992; PEARSON *et al.* 2005). However, somatic expansion has also been reported in some of these diseases. These expansions are described as dynamic mutations since they induce a continuous change of the repeat size in successive generations and throughout the life span of an affected individual. The molecular basis underlying the mechanism of repeat expansion in either germ line or somatic cells remains to be elucidated. However, many studies have shown that the expansion prone repeats form secondary structures (GACY *et al.* 1995; MITAS *et al.* 1995; GRABCZYK AND USDIN 1999; MIRKIN 2006; MIRKIN 2007; USDIN 2008), and the current thinking is that expansions are generated by the aberrant processing of these secondary structures as stated above (Figure 1.4).

Somatic expansion has been reported in patients with other Repeat Expansion Diseases and in mouse models of these diseases (TELENIUS *et al.* 1994; CHONG *et al.* 1995; UENO *et al.* 1995; MANLEY *et al.* 1999a; TANAKA *et al.* 1999; DE BIASE *et al.* 2007; VAN DEN BROEK *et al.* 2007). Data from these studies suggest that expansion of the repeat tract in the somatic cells can exacerbate disease severity and reduce the age of onset of these diseases, e.g., in the case of DM1 (MORALES *et al.* 2012) and in HD (WHEELER *et al.* 2003; SWAMI *et al.* 2009). Somatic mosaicism has been observed in human carriers of the FX PM and FM alleles (PIERETTI *et al.* 1991; PRETTO *et al.* 2014), but it was unknown whether this reflected contractions from larger alleles or expansions from smaller ones.

Somatic expansion of the FX PM alleles in mice was previously observed in our laboratory(ENTEZAM AND USDIN 2008; ENTEZAM AND USDIN 2009). However, this phenomenon was not studied extensively. Since a correlation exists between repeat number and disease symptoms as well as a direct relationship between repeat number and the likelihood of further expansion (ASHLEY-KOCH *et al.* 1998; TASSONE *et al.* 2007a; NOLIN *et al.* 2011), somatic expansion may be a clinically relevant phenomenon.

We therefore assessed somatic expansion in human cells using genomic DNA that had been previously extracted by Dr. Daman Kumari in the Usdin laboratory. We also assessed somatic expansion in different organs of male and female mice. This work showed that somatic expansions of the FX PM alleles can occur in both human cell lines and mice. We also showed that in mice somatic expansion is age-dependent, tissue-specific, and gender-biased. In addition, we have identified factors that account, at least in part, for some of these differences.

2.2 Results:

2.2.1 Expansion is seen in a human lymphoblastoid cell line

To investigate whether expansion can occur in human somatic cells, we used the PCR to amplify 100 ng of genomic DNA that had been isolated from lymphoblastoid cells cultured *in vitro*. We used the primer pair Frax AF and Frax F-Hex to amplify across the repeat as detailed in the Materials and Methods chapter. The PCR products were resolved on a 3130XL Genetic Analyzer and analyzed using GeneMapper 4.0 software. PCR amplification across the CGG•CCG repeat generates multiple PCR products that appear in GeneMapper profiles as a cluster of peaks differing by a single

repeat unit (Figure 2.1). Some of these peaks represent strand-slippage products that are generated when amplifying through long repeat tracts. The PCR products smaller than the original allele that do not change with genotype, age or tissue are likely to fall into this category. However, products that are larger than the main alleles primarily represent expansions since they increase with age and are absent in PCR profiles generated from mice with null mutation in genes, like *Msh2* or *Msh3*, that eliminate expansion as discussed in the next chapter and shown in Figure. 2.1.

Analysis of the human cell line revealed that the original allele (0 month) had ~118 repeats (Figure 2.2). At 24 months, the single peak seen in the original culture had been replaced with alleles that had gained ~16, ~35 or ~51 repeats resulting in alleles with ~ 136, 150, and 169 repeats respectively. At 42 months, alleles carrying 160, 177, 182, and 208 repeats were apparent. These alleles were ~45, ~62, ~77 and ~93 repeats larger than the original allele, respectively. Their amplification yield was low compared to the yield of PCR product from alleles detected at 24 months. This is likely due to some combination of the increased heterogeneity of alleles in the population, together with the fact that alleles with larger repeat numbers are more difficult to amplify. Thus, expansion in this cell line seemed to have occurred in a series of successive jumps that over a period of 3+ years resulted in the conversion of a PM allele to larger alleles that included some in the FM range.

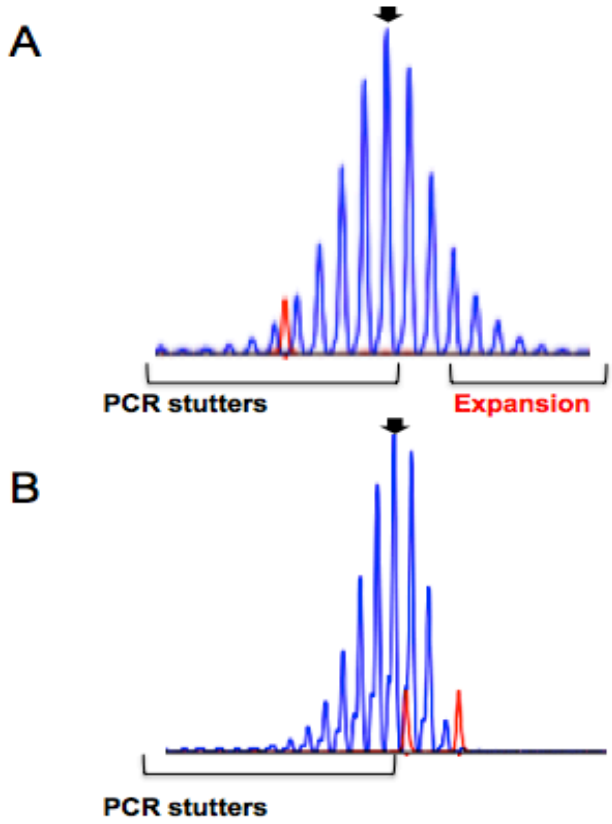


Figure 2.1 Illustration of the CGG•CCG repeat PCR product profiles in WT and *Msh2*$^{-/-}$ mice. GeneMapper Scan showing a PCR profile of a FX PM allele in an *Msh2*-wildtype background (A) and in a *Msh2* null background (B).

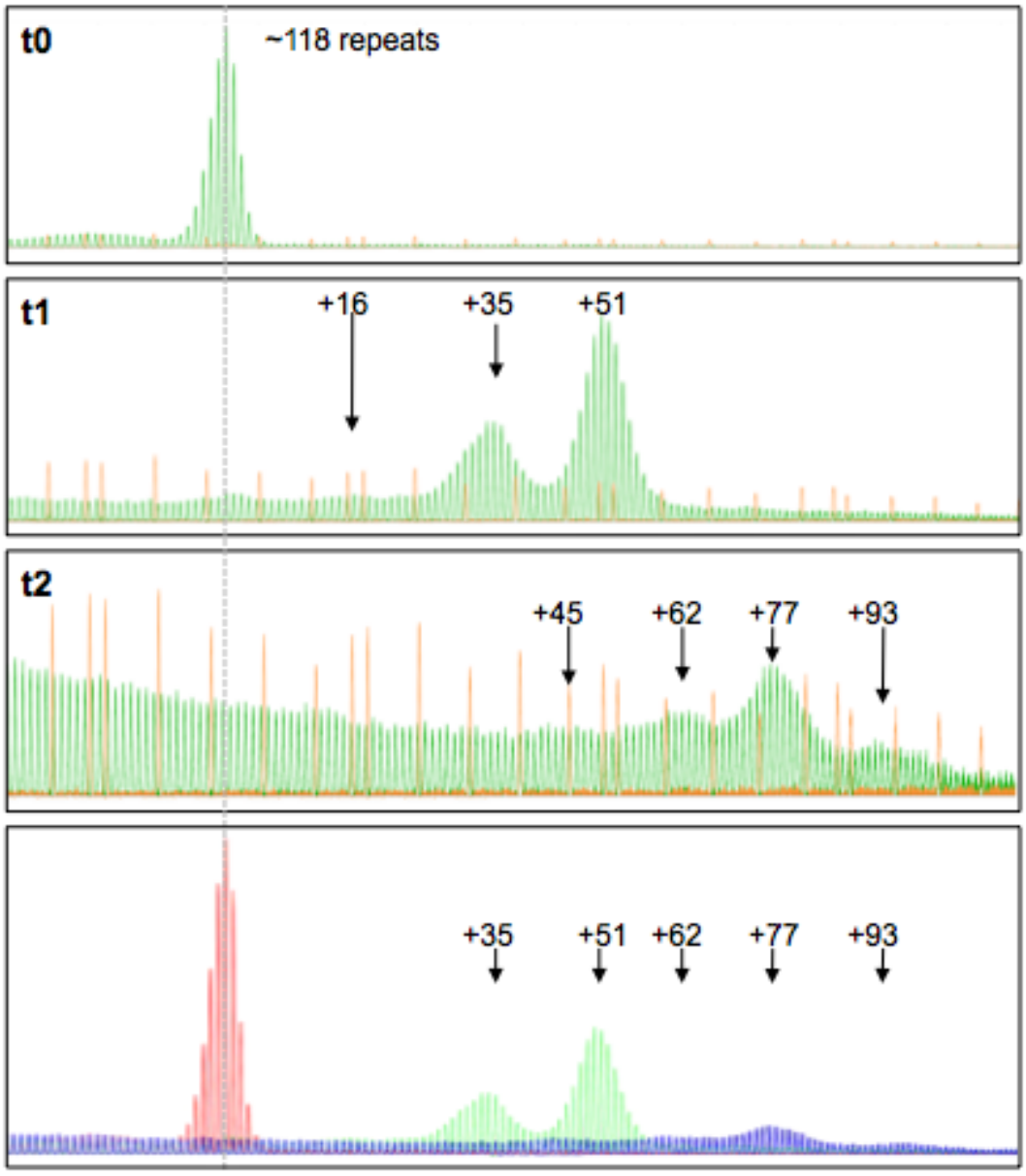


Figure 2.2 Expansions in a human PM lymphoblastoid cell line. GeneMapper scans showing *in vivo* expansion of FX PM repeats of the lymphoblastoid cell line after recurrent growth in cell culture for about 3 years. t represents different time points: t0 = 0 month, t1= 24 months, and t2= 42 months. A LIZ1200 standard was included (orange peaks) in all samples. The lower panel shows all 3-time points superimposed on one another. The red peak represents the original allele at time zero and the green and blue peaks correspond to the 24 and 42-months time points. The dotted vertical line indicates the size of the original allele corresponding at t0.

2.2.2 The extent of somatic expansion varies in different organs of male mice.

To investigate somatic instability in FX PM mice, we isolated genomic DNA from different organs of male mice at different ages. The initial repeat number in these animals was ~140. The genomic DNA was amplified by fluorescent PCR, resolved on a 3130XL Genetic Analyzer and analyzed using GeneMapper 4.0 software, as detailed in the Materials and Methods chapter. Somatic expansion was examined by comparing the repeat PCR profile in animals 2, 4 and 12 months old to the repeat profile seen in the tail DNA of the same animals at three weeks of age (tail 1). We also quantified the extent of somatic expansion using the somatic instability index (SII), a quantitative measure of the level of somatic instability developed in the Wheeler laboratory (LEE *et al.* 2010).

In the 2-month-old animals the PCR profiles of liver and testis showed a small shift to the right compared to the other organs, consistent with a low level of expansion (Figure 2.3.A). At 4 months of age, a moderate level of somatic expansion is seen in the brain, liver, and testis as evidenced by the further right shift of the PCR profile (Figure 2.3 A) and the positive SII values (Figure 2.3 B). Extensive level of somatic expansion is seen at 12 months of age (figure 2.3AB). Two widely separated peaks were seen in some organs like liver and testis, one corresponding to the original inherited allele, and another corresponding to a significantly larger allele. In other organs (e.g., brain), the increase in size of the expanded alleles was more modest. These two different types of

expansion profiles would be consistent with the idea that expansions can involve small "jumps" or large "jumps" depending on the organ involved

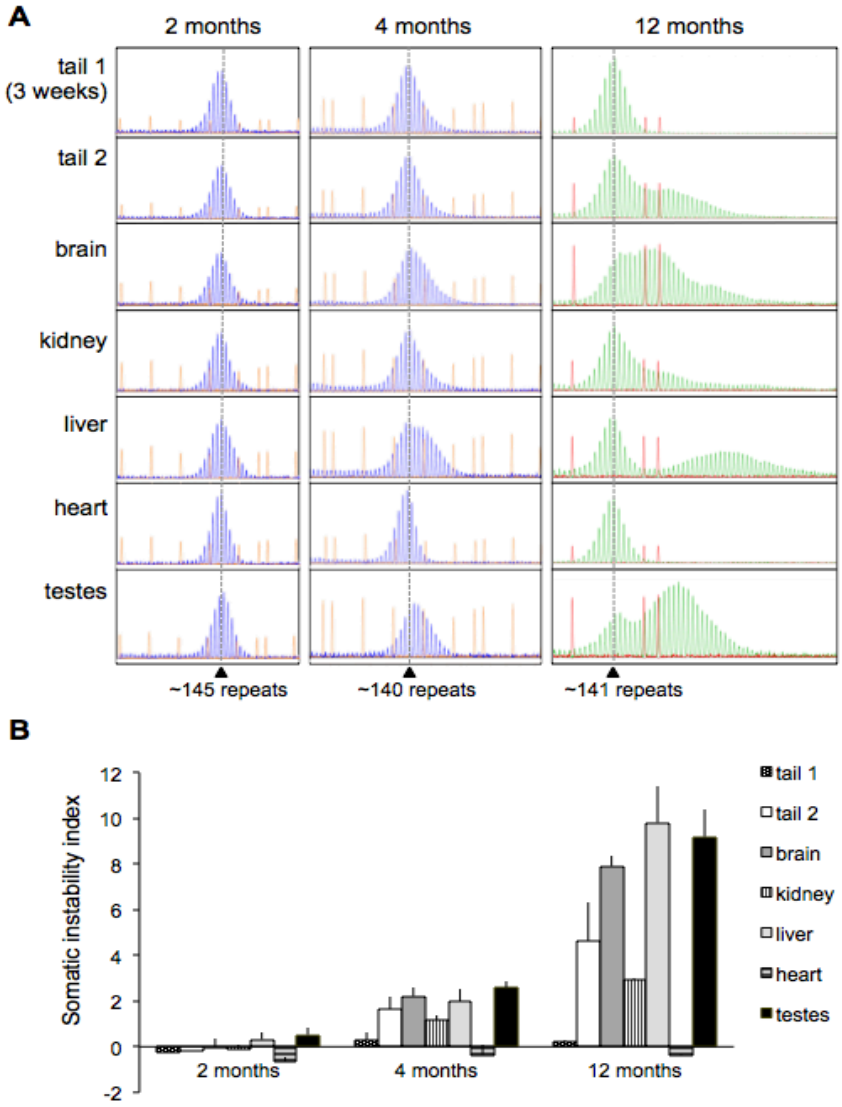


Figure 2.3 Expansions in different organs of male mice. A) Scans showing repeat expansion profile in organs of mice at 2, 4,and 12 months of age. The inherited parental allele based on the repeat number in the tail DNA taken at weaning age (tail 1). Tail 2 refers to tail taken after euthanasia. The repeat DNA profiles were generated using a Fam (blue) or Hex labeled primer (green). LIZ600 or LIZ1200 (orange) and Rox 500 molecular weight markers were used. B) Graph indicating the average somatic instability index (SII) in different organs of male mice at 2 (n=5), 4 (n=3) and 12 months (n=5). These mice had between 141 and145 repeats.

2.2.3 Somatic expansion is less extensive in female mice

To investigate whether the extent of somatic expansion is similar in female and male mice, we extracted genomic DNA from different organs of 8 female and 8 male FX PM mice. These mice were age-matched (12 months old) and all carried ~ 148 repeats. We then examined the PCR profile in different organs as described in the Material and Methods chapter. We found that while the same tissues were prone to expand as in males, expansion was much less extensive in females (Figure 2.4 A and B). Even in 24-month-old females less expansion was seen than in 12-month-old males (Figure 2.4 A and C).

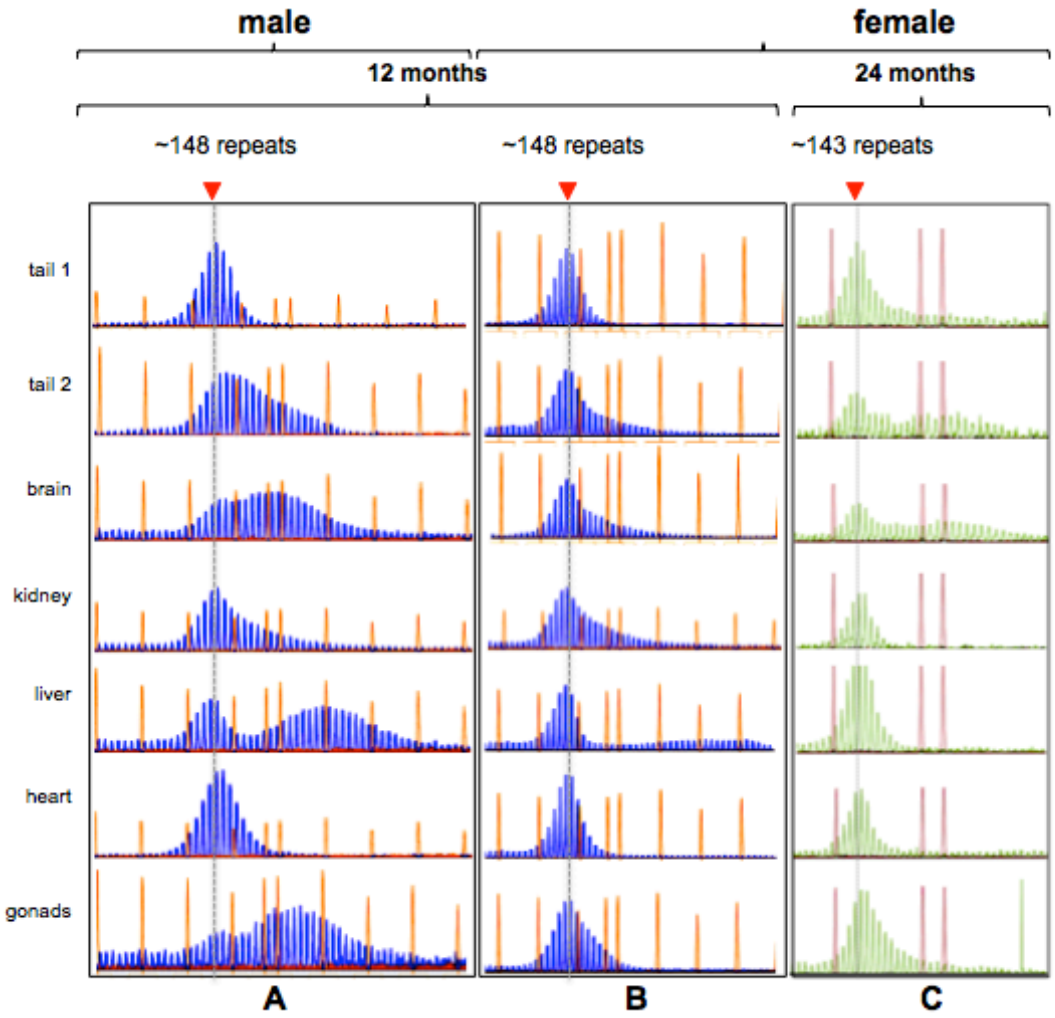


Figure 2.4 Gender bias of somatic expansions in the FX PM mice. A) Representative GeneMapper scans showing the expansion status in the indicated organs of a 12-month old male mouse. B) GeneMapper scans showing expansion profile in designated organs of a female mouse of same age and repeat number. C). GeneMapper scan showing the extent of expansion in a 24-month-old female carrying 143 repeats. The red arrow and the dotted lines show the mobility of the original inherited allele. Either LIZ600/1200™ (orange) or ROX 500™ (red) size standards were used; the GeneMapper profile was not affected in anyway by the choice of the marker used. While tail 1 indicates the tail sample cut at 3 weeks, tail 2 refers to the tail sample collected at the time of death (12 or 24 months).

The difference in the level of expansion in females versus males can perhaps be seen more clearly when PCR profiles from age and repeat-matched animals are overlaid (Figure 2.5A). We also measured the Somatic Instability Index (SII) in females and males. The SII in males was 2.9 fold higher in brain, 4.6-fold higher in liver and 6.8-fold higher in gonads than it was in females (Figure 2.5B).

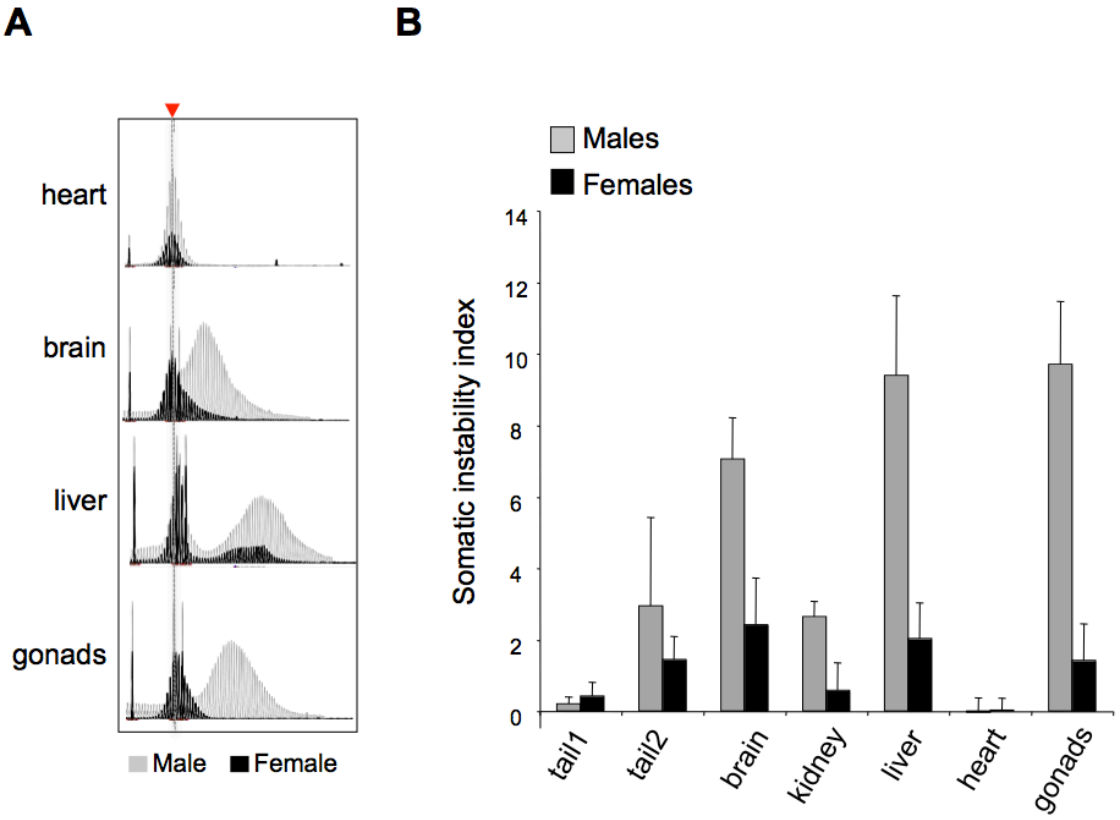


Figure 2.5 Gender bias of somatic expansions in the FX PM mice. A) Superimposed GeneMapper scans from the heart and indicated expansion prone organs of a 12 month old male and a female mouse. Both animals carried ~146 repeats. The red arrowhead and the dotted line indicate the mobility of the inherited parental allele. B) Average somatic instability index in indicated organs of one year old males (n=8) and females (n=8) carrying ~145 repeats. Tail 1 indicates the tail sample taken at 3 weeks old, at weaning; while tail 2 refers to the tail DNA sample collected at the time of death (12 months).

Since males and females show differences in the extent of expansion but are similar in terms of what organs are expansion-prone, it suggests the existence of both tissue and gender-specific factors that contribute to expansion risk.

2.2.3.1 Factors contributing to FX PM repeat expansion

2.2.3.1.1 XCI contributes to the reduced SII seen in FX PM female mice

XCI is the process by which one of the two copies of the X chromosome present in female mammals is silenced (LYON 1961). As the *Fmr1* gene is located on the X chromosome, it is possible that some of the gender differences seen in the FX PM mouse could be due to the fact that the PM allele will be on the inactive X chromosome in approximately half of cells in females heterozygous for PM. To test this hypothesis, we developed an assay that allowed us to specifically look at the PCR profile of inactive alleles. This assay took advantage of the fact that a methylation-sensitive restriction enzyme, *Sau96I*, contains a recognition site in one of the primers (Frax m5) used for the repeat PCR.

Genomic DNA from brain and gonads of female and male PM mice was either undigested or digested with *Sau96I* as described in the Material and Methods section. Since males have only one X chromosome, this chromosome is active and hence has an unmethylated *Sau96I* site. Therefore, complete *Sau96I* digestion will prevent amplification of the repeat in males. In females, ~50% of the X chromosomes will remain after digestion. Following DNA digestion, ~100 ng of undigested and digested DNA was used in the FX PM PCR assay as described in the Materials and Methods

chapter. The PCR products were resolved by capillary electrophoresis (CE) on an ABI Genetic Analyzer and the PCR profiles analyzed using GeneMapper 4.0 Software. The GeneMapper scan of the undigested female sample represents a profile of the repeats that are located on the active and inactive X chromosomes, whereas the PCR profile generated after digestion of the female sample with *Sau96I* represents the repeats located on the inactive X chromosomes only.

No PCR amplification of the PM allele was seen in the male DNA samples digested with *Sau96I* (Figure 2.6), congruent with the observation that in males there is only one X chromosome and this chromosome does not undergo XCI. However, in females a PCR profile could be seen in both undigested and digested DNA. This is consistent with the fact that the PM allele is found on the inactive X chromosome in approximately half of cells of heterozygous females and is therefore resistant to *Sau96I* 50% of the time. Expansions were seen in the undigested DNA as shown by the presence of PCR products larger than the original allele (Figure 2.6). However, there was no expansion observed in the digested samples. Thus, repeat expansion in tissues of female PM mice is seen exclusively when the PM allele is located on the active X chromosome. These data suggest a requirement for transcription or the presence of the PM allele in a transcriptionally competent region of the genome in order for expansion to occur.

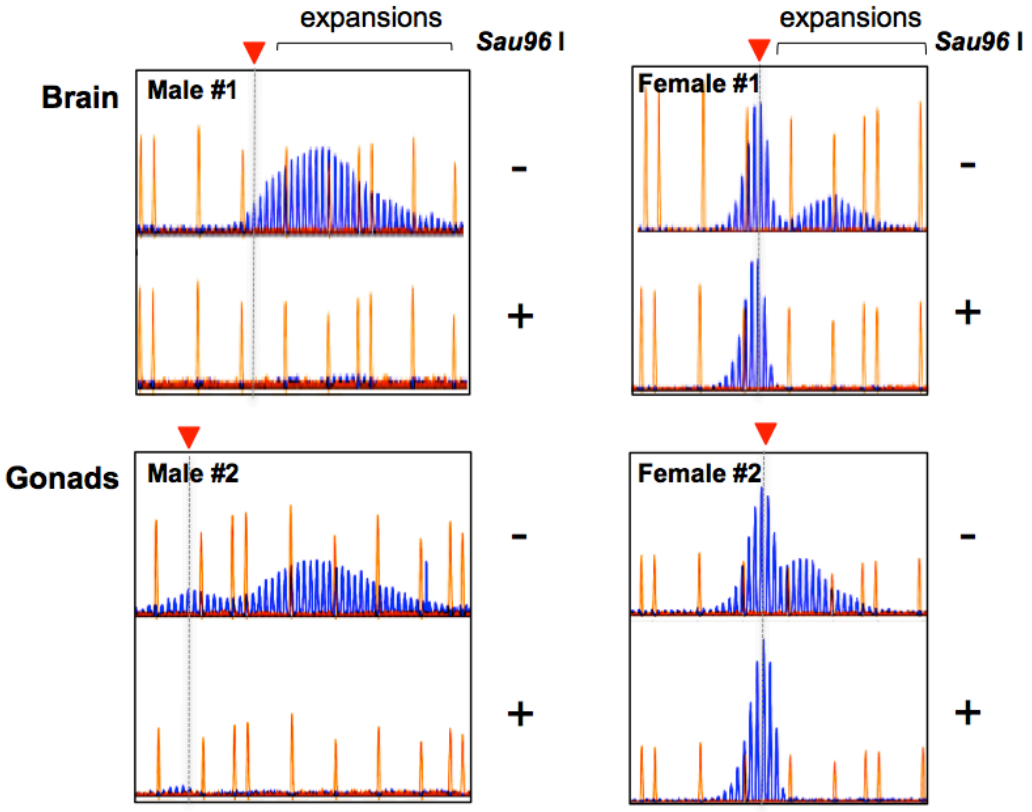


Figure 2.6 Somatic expansions on active and inactive X chromosomes. Representative GeneMapper profiles showing expansion in two males (1 brain, 1 testis) and two females (1 brain, 1 ovary) without (-) and with (+) digestion with *Sau96I*. The parental inherited allele is indicated by both the dotted line and the red arrowhead

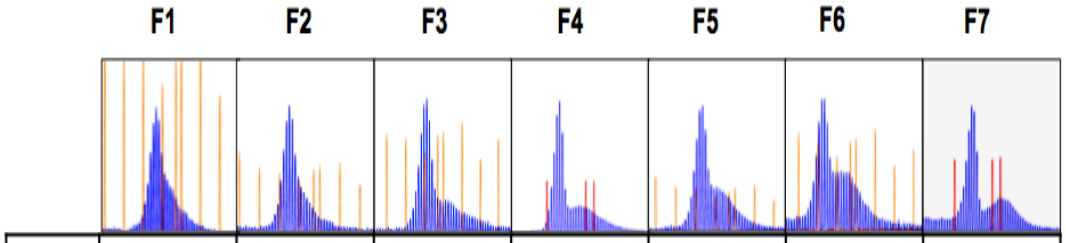
2.2.3.1.2 Skewed XCI does not explain differences in female SII.

While our data indicate that XCI contributes to the gender biases in the degree of expansion, the >2-fold differences in the SII between males and females (Figure 2.5B) cannot be completely explained by this factor alone, unless there was significant skewing of XCI. To test this idea, we examined X chromosome inactivation ratios in the brain tissues of 7 different PM females with SII ranging from 1.3 to 3.1. These SII values span the range of SII values seen in the brain of female mice.

To do this, we used a modification of the assay we used previously to examine XCI. In this assay, genomic DNA from the brain samples of female mice were digested with or without *EaeI,* a methylation-sensitive restriction enzyme that possesses a recognition site in the neomycin gene that is located downstream of the repeat in intron 1 (ENTEZAM *et al.* 2007). This gene was inserted along with the premutation repeats when the mouse line was originally made. Following digestion, the undigested and digested samples were amplified using a real-time PCR assay, as explained in the Materials and Methods chapter. The ratio of the yield of the digested DNA amplicons to the yield of the same amount of undigested material was used to determine the X inactivation ratio.

Three out of 7 females showed an X inactivation ratio of ~50:50 (active X: inactive X), 2 out of 7 showed a ~30:70 ratio, 1 female showed ~20:80 ratio, and 1 showed an 80:20 ratio. On average, the PM allele was found to be located on the active X chromosome ~ 44% of the time. We did not find a good correlation between whether the PM allele was located on the active chromosome and the level of somatic expansion

(Figure 2.7). For example, the three mice with the average SII of ~ 3.0 (F5, F6, and F7) in the brain, had the PM allele located on the active X chromosome ~32% of the time in one case, ~49% of the time in the second, and 80% of time in the third case. Similarly, the two brain samples with the lowest SII (1.3, F1 and F2) had the PM allele located on the active X chromosome, 45% and 20% of the time. These data indicate that X inactivation does not completely account for the gender differences in the levels of somatic expansion that is seen in the FX PM mice. Thus, there must be additional gender-specific factors that contribute to expansion risk.



	F1	F2	F3	F4	F5	F6	F7
SII	1.3	1.3	1.9	2.1	2.9	3.0	3.1
XCI	45:55	20:80	30:70	55:45	32:68	80:20	49:51

Figure 2.7 SII and XCI ratios in FX PM female mice. GeneMapper profile and Somatic instability index (SII) showing the degree of expansion in brain relative to XCI in one-year-old female mice. The XCI ratio is shown as the fraction of the PM allele that is located on the active X versus the fraction of the PM alleles located on the inactive X chromosome. F= female

2.2.3.1.3 Some MMR and BER proteins correlate with SII in different organs.

A number of proteins implicated in DNA repair have been shown to be important for repeat expansion in various mouse models of the Repeat Expansion Diseases (MANLEY *et al.* 1999b; SAVOURET *et al.* 2003; KOVTUN *et al.* 2007; MOLLERSEN *et al.* 2012; LOKANGA *et al.* 2014). To test whether the levels of tissue-specific expression of these proteins could explain the propensity of some organs to expand, we examined the levels of these and other proteins in the brain, heart, liver, kidney and testis of male mice at 6 months of age by Western blot.

As can be seen from Figure 2.8A, neither the levels of MSH2 nor MSH3 correlated completely with the extent of expansion in different organs. For example, while expansion is more extensive in liver than in kidney, the level of MSH2 was higher in kidney while the levels of MSH3 are similar in these two organs. However, when MSH2, MSH3 and MSH6 are considered together, a correlation is seen. Of all the proteins tested in the BER pathway, only APE1 and FEN1 seemed to correlate with the levels of expansion in different organs (Figure 2.8B). The levels of these proteins were both highest in the testes, followed by liver, brain, kidney, and heart.

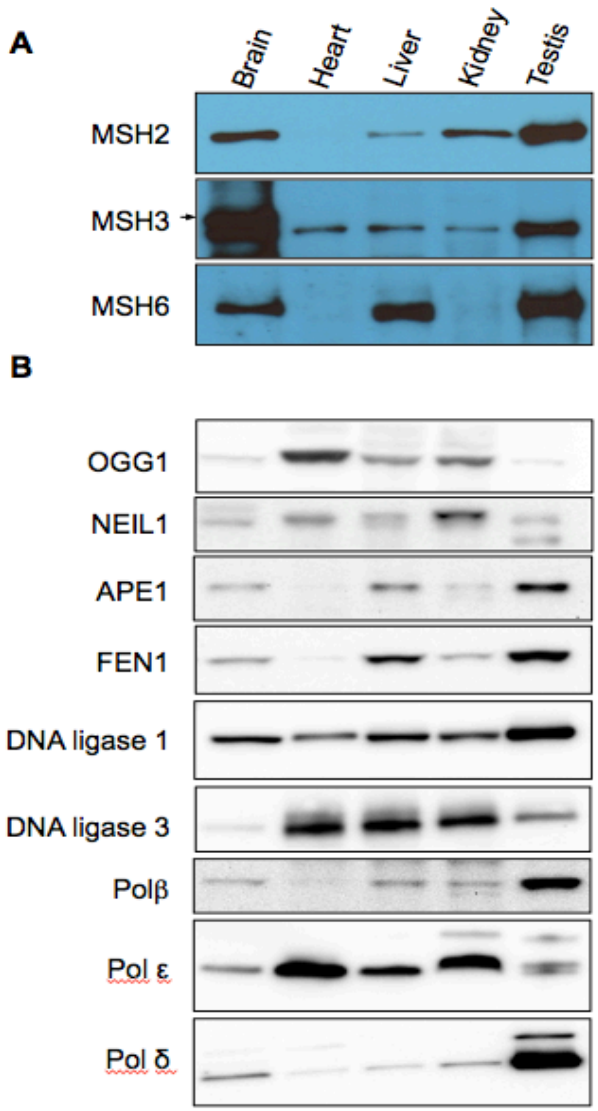


Figure 2.8 MMR and BER proteins in indicated organs of FX PM mice. Thirty micrograms (30ug) of protein extracted from the brain of one year old male FX PM mice (n=3) were resolved on a gradient Tris-acetate gel, electroblotted to nitrocellulose membrane, and probed with the indicated antibodies. A) Representative data displaying the amounts of MSH2, MSH3, MSH6 proteins in brain, heart, liver, kidney, and testis B) Levels of OGG1, NEIL1, APE1, FEN1, DNA ligase 1, DNA ligase 3, Polβ, Polε, and Polδ . Equal amount of protein from the organs of each animal were pooled and protein detected by Western blot.

.

2.2.3.1.4 Lack of gender biases in the levels of DNA repair proteins in FX PM mice

To test whether there are differences in the expression of various DNA repair proteins that could explain the gender differences in the level of expansion seen in our mouse, we examined in males and females the levels of expression of a number of autosomal DNA repair genes previously shown to be involved in expansion in various models of some Repeat Expansion Disorders. We also compared the expression levels of various DNA repair genes that are located on X chromosome in male and female mice. We used Western blotting to compare the amounts of the MMR proteins MSH2, MSH3, and MSH6, and the BER proteins OGG1 and NEIL1 in the liver tissues of 4 male and 4 female mice. We also measured the protein levels of UBE1, expressed by *Ube1* (ubiquitin-like modifier activating enzyme 1), an X-linked gene that escapes X-inactivation in mice. *Ube1* is a DNA repair gene that catalyzes the first step in ubiquitin conjugation to mark cellular proteins for degradation (KUDO *et al.* 1991). We also measured the levels of SMC1a, a protein expressed by *Smc1a* (structural maintenance of chromosomes 1A), another X-linked gene that escapes X inactivation in mice (YANG *et al.* 2010). In all cases, β actin was used as a loading control.

No significant difference in the expression of MSH2, MSH3, MSH6, OGG1 and NEIL1 proteins were seen in organs of female and male FX PM mice (Figure 2.9). The amount of the X-linked DNA repair proteins, UBE1 and SMC1a, were also similar in both males and females. Thus, there must be other contributing factors that reduce the risk of expansion in female mice.

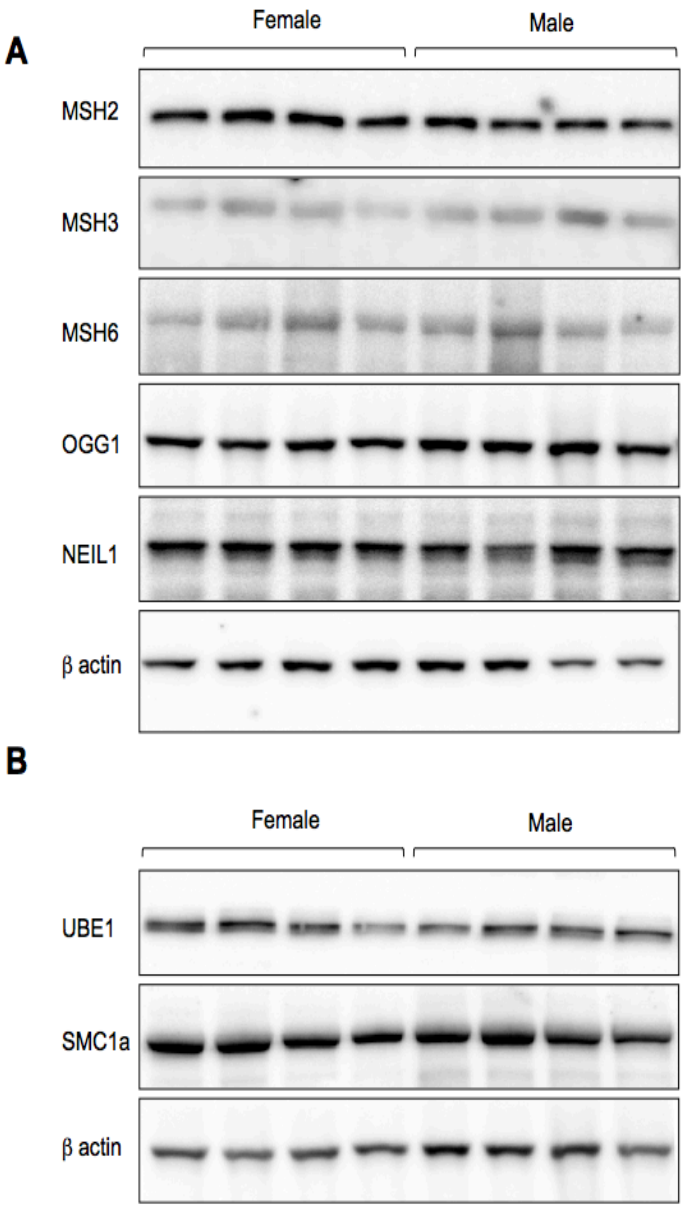


Figure 2.9 DNA repair protein levels in female and male FX PM mice. Thirty micrograms of protein extracted from the liver of one year-old male (n=4) and female (n=4) FX PM mice were resolved by electrophoresis on a gradient Tris-acetate gel, electroblotted onto nitrocellulose membrane, and probed with the indicated antibodies. β actin was used as loading control. A) Amounts of DNA repair proteins implicated in one or more mouse models of the Repeat Expansion Diseases. B) Levels of DNA repair genes located on the X chromosome.

2.2.3.1.5 Lack of protective effect of estrogen on expansion in FX PM females mice

Estrogen has been shown to protect against oxidative damage (RICHARDSON *et al.* 2012). Since oxidative damage has been shown to increase expansion risk in the FX PM mice, we hypothesized that estrogen may decrease the risk of expansion in females. To test this idea, we ovariectomized 3 PM females at 16 days old and compared the extent of somatic expansion in the tail, brain, kidney, liver, heart, and ovary of these mice at 12 months of age to unovariectomized mice of the same age by PCR as described in the Material and Methods chapter. We found no differences in any organ in the extent of somatic expansion between ovariectomized and unoveriectomized FX PM mice (Figure. 2.10). These data suggest that a protective effect of estrogen does not explain the lower level of somatic expansion in females.

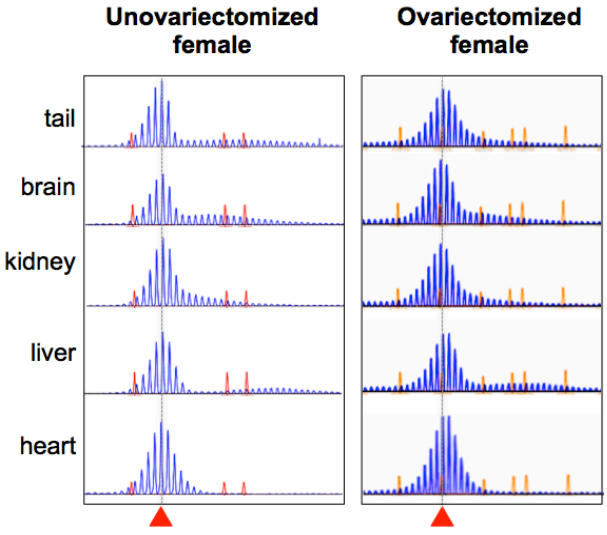


Figure 2.10 Somatic expansion in ovariectomized PM FX female mice. Representative Gene Mapper profiles of one-year-old unovariectomized and ovariectomized mice with similar repeat numbers. The dotted line and the red arrowhead indicate the parental inherited allele.

2.3 Discussion

We have shown that expansion can occur in human somatic cells. We observed the stepwise conversion of a PM allele with ~118 repeats to one with 208 repeats i.e., in the FM range (Figure 2.2). We have also shown that the degree of somatic expansion of the FX PM alleles in different organs of mice is age-dependent, tissue-specific, and gender-biased (Figures 2.3, 2.4, and 2.5). Somatic expansion of the FX PM alleles in these animals becomes clearly apparent at 4 months of age and increases with age. This would be consistent with a mechanism in which somatic expansions accrue by incremental increases in repeat number over a period of time, as shown by previously published data on somatic expansion in mouse models of other Repeat Expansion Diseases (ISHIGURO *et al.* 2001; VAN DEN BROEK *et al.* 2002; GONITEL *et al.* 2008).

In the FX PM mouse model organs such as brain, liver, and gonads were more prone to expand than kidney, with the PM allele being stable in heart. The highest level of expansion is seen in the brain and liver, where most of the cells are post-mitotic. This supports the hypothesis that expansion is not dependent on chromosomal replication. This is consistent with work by others in the Usdin laboratory showing that expansion occurs in oocytes, which do not divide. It is also consistent with an observed maternal age-effect on the expansion risk in humans (YRIGOLLEN *et al.* 2012). Thus, our data lends support to the idea that expansion in FX PM mouse model, and in human PM carriers, may be more likely to result from aberrant DNA damage repair than events occurring during normal genomic replication.

The strong tissue-specificity of expansion that is seen in the FX PM mouse, as well as in a number of mouse models of other Repeat Expansion Diseases, has led to the suggestion that tissue-specific factors may be involved (LIA *et al.* 1998; FORTUNE *et al.* 2000). However, not all organs expand to the same extent in mouse models of different Repeat Expansion Diseases. For example, while organs like the kidney show relatively little expansion in the FX PM mouse model, this organ shows extensive expansion in other mouse models (FORTUNE *et al.* 2000; VAN DEN BROEK *et al.* 2007).

Variations in the levels of DNA repair proteins have been suggested to account for the tissue-specificity of repeat expansion in mouse models of HD and DM1(GOULA *et al.* 2009; TOME *et al.* 2013; MASON *et al.* 2014). For example, in HD it has been reported that the levels of BER and MMR proteins correlate with increased somatic CAG instability (GOULA *et al.* 2009; SERIOLA *et al.* 2011). However, the levels of these proteins were examined only in a few sub-regions of the brain or a few cell lines. When we examined a wider range of different organs of the FX PM mouse and compared the levels of a much larger set of proteins, no good correlation between the extent of expansion and the levels of any of the previously identified proteins. However, while neither the level of MSH2 or the level of MSH3 correlated well with the extent of expansion in all organs, the combination of MSH2, MSH3 and MSH6 did have a correlation (Figure 2.8A). In addition, the levels of APE1 and FEN1 also showed a correlation (Figure 2.8B).

We also showed that somatic expansion of the FX PM is more extensive in males than females, while the organs prone to expand remained the same in both

genders (Figures 2.4 and 2.5). We found that this gender difference was due in part to the fact that *Fmr1* is an X-linked gene and that expansion occurred only on a PM allele that was located on the active X chromosome (Figure 2.6). This observation suggests that transcriptional activity of the gene or an open chromatin configuration is necessary for the expansion to occur. These data support previous observation that in human carriers of methylated FM alleles, the repeat is stable (GLASER *et al.* 1999). It also lends support to previous suggestions that transcription plays a role in repeat expansion in mouse models of HD and DM1(LIA *et al.* 1998; GOULA *et al.* 2012). However, since chromosomal context affects expansion and the chromosomal context of the transgenes being compared in these cases differed from one another, it was not possible to definitively implicate transcription/open chromatin in expansion based on these experiments. But, since *Fmr1* is an X-linked gene, we were able to compare simultaneously the extent of expansion of the same PM allele in the same sequence context on the active and inactive X chromosome in the same animal. Thus, our data allowed us to definitively demonstrate a requirement for transcription and/or an open chromatin configuration in the FX expansion in mice. Our data are also consistent with a retrospective examination of data from human females who carry the PM allele (GRASSO *et al.* 2014).

However, as was shown by others in the Usdin laboratory, there is no good correlation between the level of *Fmr1* expression in expansion prone organs and the degree of somatic expansion in the same organs (LOKANGA *et al.* 2013). Furthermore, there was no gender bias in the transcription of the *Fmr1* gene that could explain the gender differences on somatic expansions that is seen in this mouse model (LOKANGA *et*

al. 2013; ZHAO AND USDIN 2014). Therefore, there must be additional unknown factors that are necessary for expansion.

We have analyzed a number of additional potential factors that could contribute to the gender differences on somatic expansion in the FX PM mouse, including the effect of estrogen and differences in the expression of a number of DNA repair genes known to be important for expansion (Figures 2.9 and 2.10). However, none of these factors are likely contributors to the gender difference seen on somatic expansions in our mouse model either. Data generated by others in the Usdin laboratory have shown that many genes involved in the removal of damaging reactive oxygen species (ROS) are highly expressed in female than male FX PM mice. Since we know that oxidative stress increases expansion risk (ENTEZAM *et al.* 2010), it is possible that these gender differences in the sensitivity to oxidative stress contribute to the gender biases on somatic expansion that we have observed.

In summary, we have shown that somatic expansion of FX PM alleles occurs in both human cells and in mice. In mice, somatic expansion is age-dependent, variable in different tissues, and occurs more frequently in males than females. Since disease symptoms are related to repeat number in humans, it is possible that somatic expansion contributes to disease risk and that identifying factors that reduce somatic expansion can be clinically useful. Furthermore, since somatic expansion is less frequent in PM female mice, it might also suggest that human females who are carriers of PM alleles may have lower risk for somatic expansion than human males.

Chapter 3: Effect of a null mutation in the MMR genes, *Msh2* and *Msh6,* on the FX PM repeat expansion

3.1 Introduction

The Fragile X related disorders are caused by the expansion of a CGG•CCG repeat tract in the 5' UTR of the *FMR1* gene. We have shown in the previous chapter that somatic expansion of the FX PM alleles occurs in human cells and in the FX PM mouse and that this expansion is tissue-specific with brain, liver and testes being more prone to expand than the kidney and heart. However, the molecular basis of FX PM repeat expansion and the factors that can explain the tissue specificity of expansion in the FX PM mouse model are not well understood.

The role of the mismatch repair (MMR) system in repeat expansion has been studied in a number of mouse models of various Repeat Expansion Disorders. However, the effects of MMR deficiency on the frequency of somatic and/or germ-line repeat expansions in these diseases were variable. For instance, in the mouse model of DM1, null mutations in *Msh2* gene, the key enzyme of the MMR pathway, resulted in suppression of somatic expansion and of most paternally and maternally transmitted expansions (SAVOURET *et al.* 2003). Furthermore, an *Msh2* null mutation was also shown to prevent somatic expansion (MANLEY *et al.* 1999a) and to abolish paternally transmitted but not maternally transmitted expansions in the mouse model of HD (WHEELER *et al.* 2003). However, in a transgenic mouse model of FRDA, an *Msh2* null mutation failed to reduce intergenerational expansion (EZZATIZADEH *et al.* 2012).

Studies of MSH3 and MSH6, the two MSH2-binding partners in mammalian cells, also resulted in variable outcomes. In a DM1 knock-in mouse model, an *Msh3* null mutation abolished somatic expansions whereas an *Msh6* null mutation increased somatic expansions (VAN DEN BROEK *et al.* 2002). In a transgenic mouse model of this same disease, MSH3, but not MSH6, was shown to be involved in germ line repeat expansion (FOIRY *et al.* 2006). In a mouse model of HD, a null mutation of *Msh3* did not have significant effect on intergenerational repeat expansion, but eliminated somatic expansions (DRAGILEVA *et al.* 2009). While MSH6 was shown not to be involved in the repeat expansions, it was shown instead to protected against parental intergenerational repeat contractions (DRAGILEVA *et al.* 2009). In contrast, in a transgenic mouse model of FRDA MSH6 was shown to be involved in the formation of GAA repeat somatic expansions (BOURN *et al.* 2012). Additionally, work in FRDA patient cells also supports a role for MSH6 in somatic expansions (DU *et al.* 2012). Thus, FRDA is the only Repeat Expansion Disease for which published data implicates MSH6 in the generation of repeat expansions.

Thus, the effects of the mismatch repair proteins on repeat instability appears to differ in different Repeat Expansion Diseases and in some cases to vary between germ line and somatic cells. This may indicate that different mechanisms are involved in expansion in different diseases and in different cell types. In an effort to better understand the mechanism of repeat expansion in the FXDs as well as factors that can explain the tissue specificity seen in the FX PM mouse model, we investigated the potential role of MSH2 and MSH6 in repeat expansion in the FXD mouse model. It

should be noted that the role of MSH3 in this mouse has been studied (ZHAO *et al.* 2015).

3.2 Results:

3.2.1 Somatic expansion of the FX PM alleles is reduced in MSH2 or MSH6 deficient mice.

To investigate whether the MSH2 and MSH6 proteins promote or prevent somatic expansion in FX PM mice, we analyzed somatic instability in multiple organs from WT mice of same age, gender and repeat number that were heterozygous or nullizygous for either *Msh2* or *Msh6*. We first crossed the FX PM mice to mice that have a null mutation in either *Msh2* or *Msh6* in order to generate mice that have the PM allele and that are heterozygous or nullizygous for either *Msh2* or *Msh6*. Since *Msh2* and *Msh6* null mutations predispose these mice to cancer and thus these animals die young, we primarily assessed somatic instability in mice that were less than 12 months old. However, 3 *Msh6* null female mice survived to twelve months of age and somatic instability was assessed in these animals as well.

We isolated genomic DNA from tail, brain, kidney, liver, heart, and gonads of mice for somatic instability analysis in *Msh2* and *Msh6* mutant mice as described in the Materials and Methods chapter. With the exception of the 12-month-old *Msh6*$^{-/-}$ mice where only 3 animals were available, we analyzed 6 animals per genotype. The genomic DNA was amplified by PCR using the primer pair Frax m5 and Frax m4, where Frax m4 was fluorescently labeled with either 6-carboxyfluorescein (FAM) or 4, 7, 2′, 4′, 5′, 7′-hexachloro-6-carboxyfluorescein (HEX). The choice of the fluorescent label does

not affect the results in any way (per our observation). The PCR product was then resolved on an ABI 3130XL Genetic Analyzer and the data analyzed using GeneMapper 4.0 software, as detailed in the Materials and Methods chapter.

The effect of *Msh2* or *Msh6* mutations on somatic expansion was determined by comparing the number of CGG•CCG repeats in the DNA from different organs at time of death to the repeat number found in the tail DNA of the same animal at three weeks of age (tail 1). We also quantitated the extent of somatic expansion using the somatic instability index (SII) as described in the Materials and Methods chapter (LEE *et al.* 2010). We then compared the degree of somatic instability in animals that were WT, heterozygous or nullizygous for either *Msh2* or *Msh6*.

3.2.1.1 Somatic expansion is eliminated in MSH2 deficient mice.

Analysis of 4 and 6 month old $Msh2^{+/+}$ (WT) mice showed evidence of somatic expansion in most organs consistent with our previous findings that somatic expansion is progressive and tissue specific (Chapter 2). In contrast, analysis of $Msh2^{-/-}$ mice showed no evidence of somatic expansion in any of the organs examined in either 4 or 6 month old mice as evidenced by the fact that the GeneMapper profiles for organs that are expansion prone in $Msh2^{+/+}$animals are indistinguishable from organs like heart that show no somatic expansion and similar to the profile generated from tail DNA taken soon after birth (Figure. 3.1A). This was consistent with the fact that all tissues had similar very low SII values at both 4 and 6 months (Figure. 3.1B).

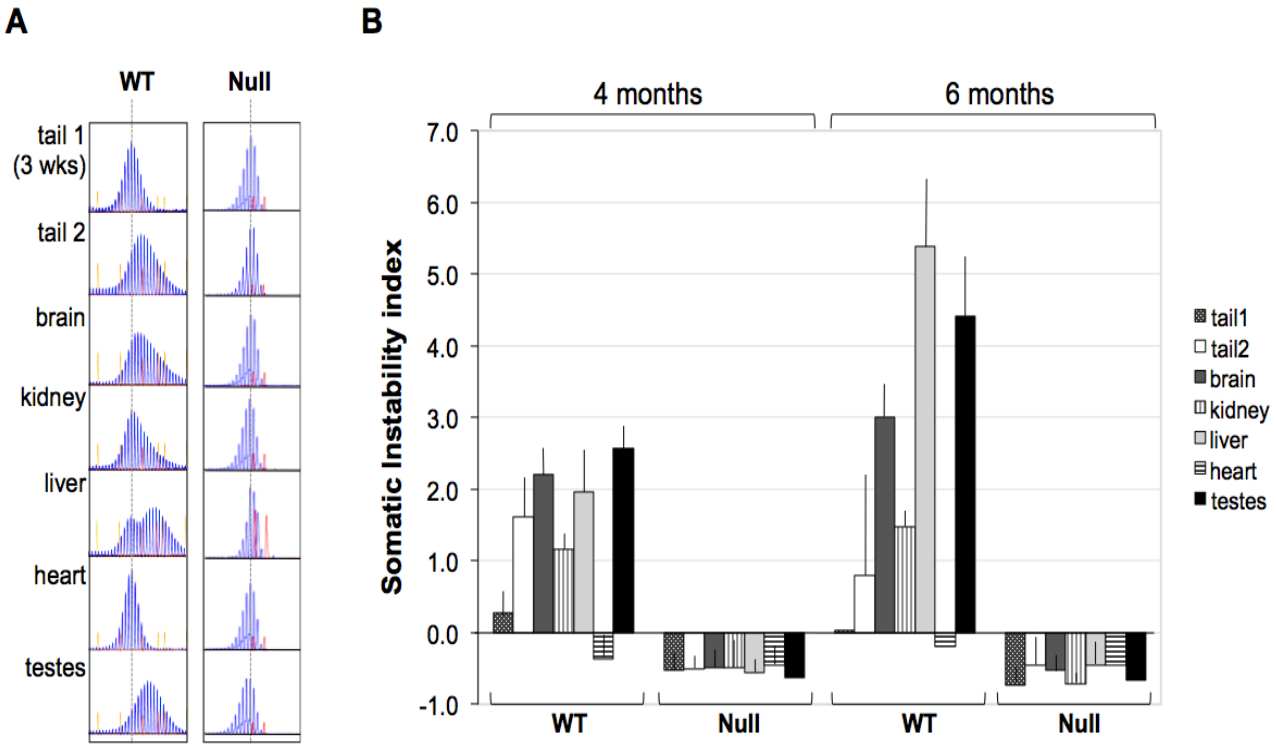


Figure 3.1 Somatic expansion in MSH2 deficient mice. (A) Representative GeneMapper scan showing expansion in indicated organs of a 6-month-old *Msh2*$^{+/+}$ and a 6-month-old *Msh2*$^{-/-}$ mouse. The dotted line indicates the original allele size in tail 1 (3 weeks). (B) Average somatic instability index (SII) in different organs of *Msh2*$^{+/+}$ (n=6) and *Msh2*$^{-/-}$ (n=6) mice at 4 and 6 months of age.

The slight negative values of the SII do not indicate an increase in somatic contraction since the SII does not change with age and it resembles at both ages the SII of heart, an organ that shows no expansions or contractions in WT animals. The low negative values of SII likely reflect the presence of stutter bands, PCR artifacts produced during PCR amplification of short tandem repeats (WALSH *et al.* 1996). Our data thus suggest that MSH2 is essential for somatic expansion in FX mice.

We also analyzed somatic expansion in brain, liver and testes of *Msh2*$^{+/-}$ mice at 12 months of age since *Msh2* heterozygous mice lived longer than *Msh2* null animals. We found that even the loss of one *Msh2* allele was sufficient to reduce significantly somatic expansion in most organs (Figure 3.2A). The average repeat number added in the brain, liver, and testis of *Msh2*$^{+/+}$ mice was 9.5, 20.5, and 12.5 respectively, and 4.2, 11.8, and 9.9 in *Msh2*$^{+/-}$ mice. While the difference in repeats size in the brain and liver of *Msh2*$^{+/+}$ versus *Msh2*$^{+/-}$ mice was statistically significant, the difference in repeat size in the testis of *Msh2*$^{+/+}$ versus *Msh2*$^{+/-}$ mice did not reach the preset statistical significance of $p<0.05$ (Figure 3.2B). Though, the SII was significantly reduced in all organs of *Msh2* null animals, no allele smaller than the original allele (tail 1 or heart) was observed in any organ and at any age. Thus, there was no evidence of increased somatic contraction in these animals.

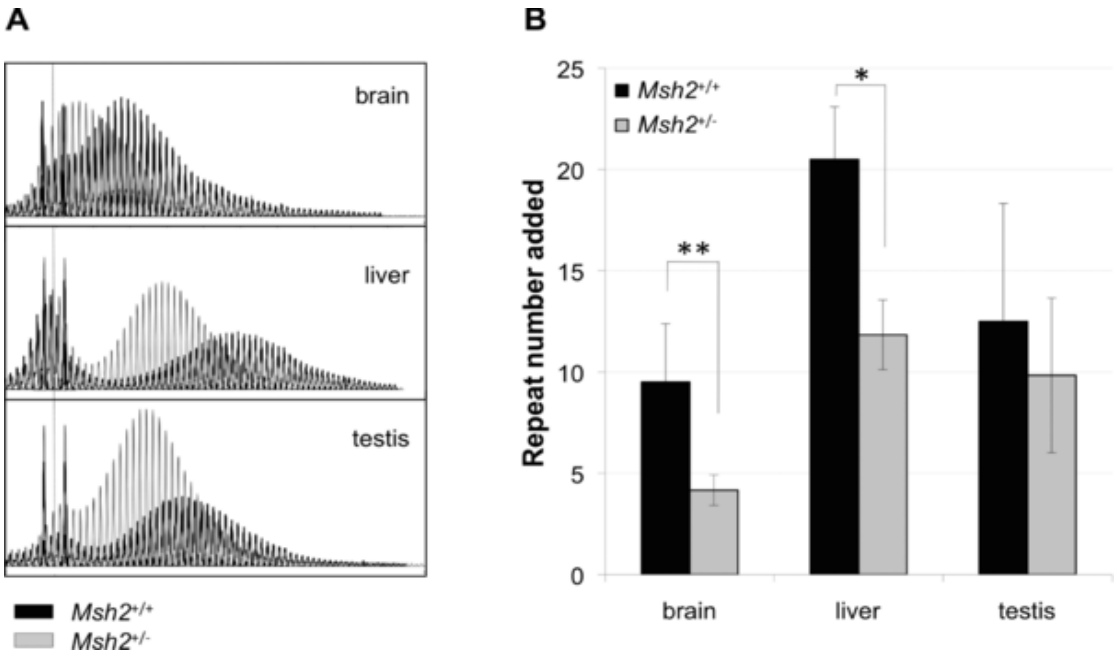


Figure 3.2 Expansion in the indicated organs of *Msh2* heterozygous FX PM mice. (A) Representative overlaid GeneMapper scans showing expansion in a one-year-old *Msh2*$^{+/+}$ and a one-year-old Msh2$^{+/-}$ mouse. The dotted line shows the parental inherited allele size in tail 1 (3 weeks). (B) The average repeat number added to the original allele in indicated organs of 6 *Msh2*$^{+/+}$ and 6 *Msh2*$^{+/-}$ male mice. The error bars indicate the standard deviation. The asterisks indicate organs with significantly different numbers of repeats added by Student's t-test (*p< 0.01 and ** p< 0.001).

3.2.1.2 Somatic expansion is reduced in MSH6 deficient mice

In mammalian cells, MSH2 forms a heterodimer with either MSH6 or MSH3 to form MutSα and MutSβ complexes respectively (KOLODNER 1996; KOLODNER AND MARSISCHKY 1999; KUNKEL AND ERIE 2005; LI 2008). In order to see whether MutSα is involved in FX PM repeat expansion, we examined the effect of the *Msh6* null mutation on somatic expansion in FX PM mice. We examined repeat expansion in somatic cells of 6 *Msh6*$^{-/-}$ male mice at 6 month of age. Since we have shown in the previous chapter that somatic expansion is age dependent and less frequent in female than in male FX PM mice, we limited somatic instability analysis in female mice to the 3 animals that survived to 12 months of age.

Analysis of the data from the six 6-month-old male mice showed that the loss of MSH6 significantly reduced the extent of somatic expansion in all expansion-prone organs (Figure 3.3). While the repeat size in heart was stable in both *Msh6*$^{+/+}$ and in *Msh6*$^{-/-}$ mice, expansion was considerably reduced in the tail, liver, and testis ($p < 0.0001$) and in brain ($p < 0.001$) of *Msh6*$^{-/-}$ mice. The SII in the tail, brain, liver, and testis were 1.1, 4.0, 3.3 and 1.6 respectively in *Msh6*$^{-/-}$ mice and 6.0, 6.0, 8.0, and 9.2 in *Msh6*$^{+/+}$ mice. Taken together these data suggest that MSH6, and thus MutSα, contributes either directly or indirectly to expansion of the FX PM alleles. However, the effect of the loss of MSH6 is not as severe as the loss of MSH2 (this chapter) or the loss of MSH3 where no expansions are seen in null animals (ZHAO *et al.* 2015).

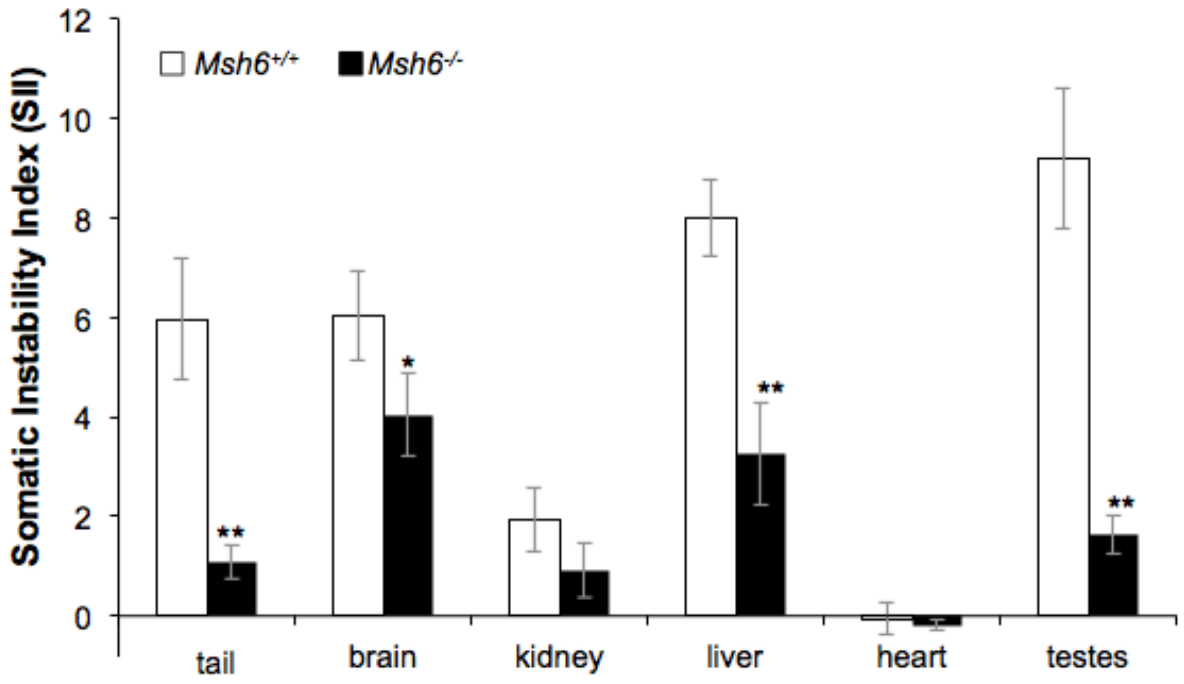


Figure 3.3 Expansion as measured by the SII in indicated organs of *Msh6*$^{+/+}$ and *Msh6*$^{-/-}$ male FX PM mice. The data represents an average of 6 *Msh6*$^{+/+}$ and 6 *Msh6*$^{-/-}$ animals at 6 months of age. The organs marked with asterisks represent those with statistically different SIIs by Student's t-test (*p < 0.01 ** p< 0.001, and *** p< 0.0001).

Analysis of GeneMapper scans of one year-old female *Msh6*$^{+/+}$ mice showed a moderate level of somatic expansion in all expansion prone organs including brain, liver and ovaries consistent with our previous data that showed that somatic expansion in female mice occurs less frequently than in males (Chapter 2). In contrast, analysis of GeneMapper profile of *Msh6*$^{-/-}$ mice revealed that the extent of somatic expansion was significantly reduced in the tail taken at time of euthanasia (p=0.001) and in the ovaries (p< 0.001) of *Msh6*$^{-/-}$ mice with the SII of 2.5 in tail and 0.59 in ovary compared to 5.9 and 2.75 in tail and ovary of *Msh6*$^{+/+}$ mice respectively (Figure 3.4). However, the

difference in the SII of the brain, liver, and kidney was not statistically significant between *Msh6*$^{+/+}$ and *Msh6*$^{-/-}$ mice. As in *Msh2*$^{-/-}$ mice, no allele smaller than the original allele was seen in any organs of either female or male *Msh6*$^{-/-}$ mice, confirming that somatic contraction does not occur or is rare in FX PM mice.

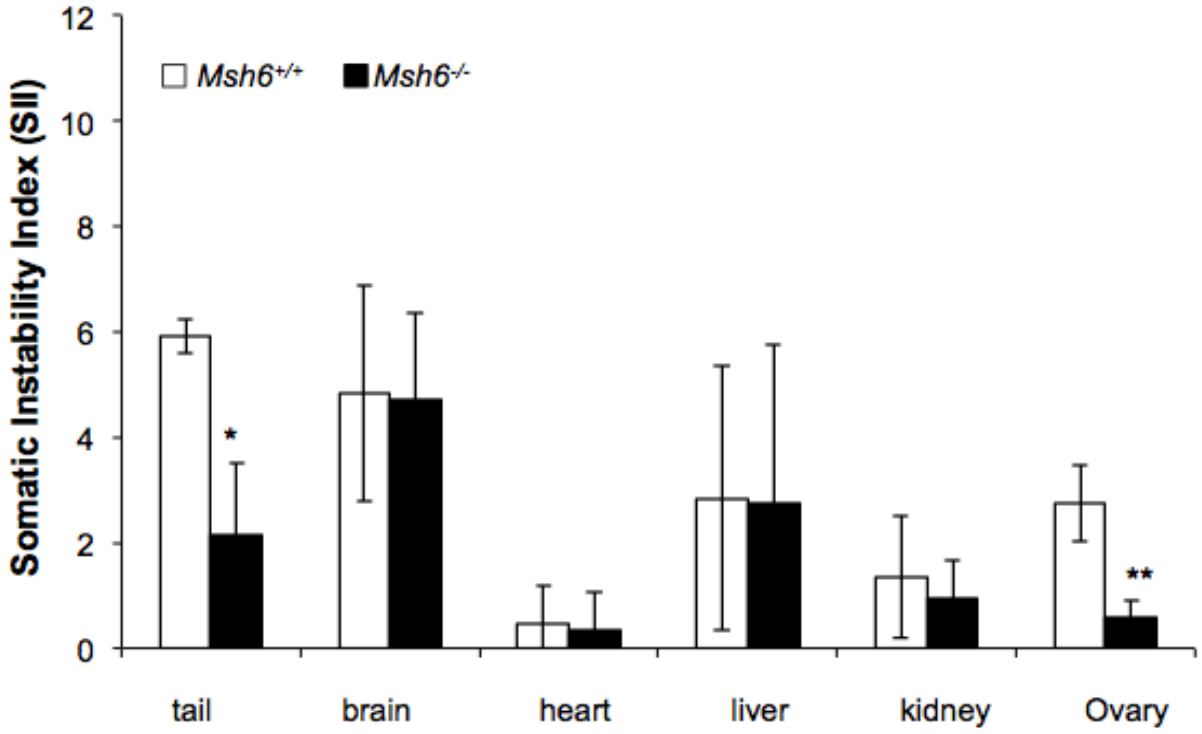


Figure 3.4 Expansion as assessed by the SII in indicated organs of *Msh6*$^{+/+}$ and *Msh6*$^{-/-}$ female FX PM mice. Four *Msh6*$^{+/+}$ and 3 *Msh6*$^{-/-}$ females were analyzed at 12 months of age. The asterisks indicate the organs in which the loss of MSH6 had a significant effect on the SII by Student's t-test (**p < 0.001 and *** p< 0.0001).

3.2.2 Germ line expansion of FX PM alleles is reduced in MSH2 and MSH6 deficient mice.

In order to evaluate the role of MSH2 and MSH6 on the intergenerational instability of the CGG•CCG repeat, we crossed mice that had one PM allele and were heterozygous or nullizygous for either *Msh2* or *Msh6* to mice that had a wild type *Fmr1* allele and the same *Msh2* or *Msh6* genotype as detailed in the Materials and Methods section. We then evaluated the frequency of expansion, contraction, and unchanged alleles in the offspring of these mice. For each genotype at least three breeding pairs were set up. We followed instability on both maternal and paternal transmission. Since *Msh2* and *Msh6* null mutations predispose these mice to cancer and thus these animals die young, we only assessed intergenerational instability in mice less than 12 months old.

Genomic DNA was extracted from tail biopsies of the animals at weaning. Amplification of the FX PM repeats and repeat size analysis were carried out as described in the Materials and Methods chapter. Intergenerational instability was determined by comparing the repeat size in the transmitting parent with that of the progeny to determine the frequency of expansion, contraction, and unchanged alleles (Figure 3.5). The parental and offspring PCR amplicons were resolved at the same time in a 3130 XL Genetic Analyzer and data analyzed using GeneMapper 4.0 software

.

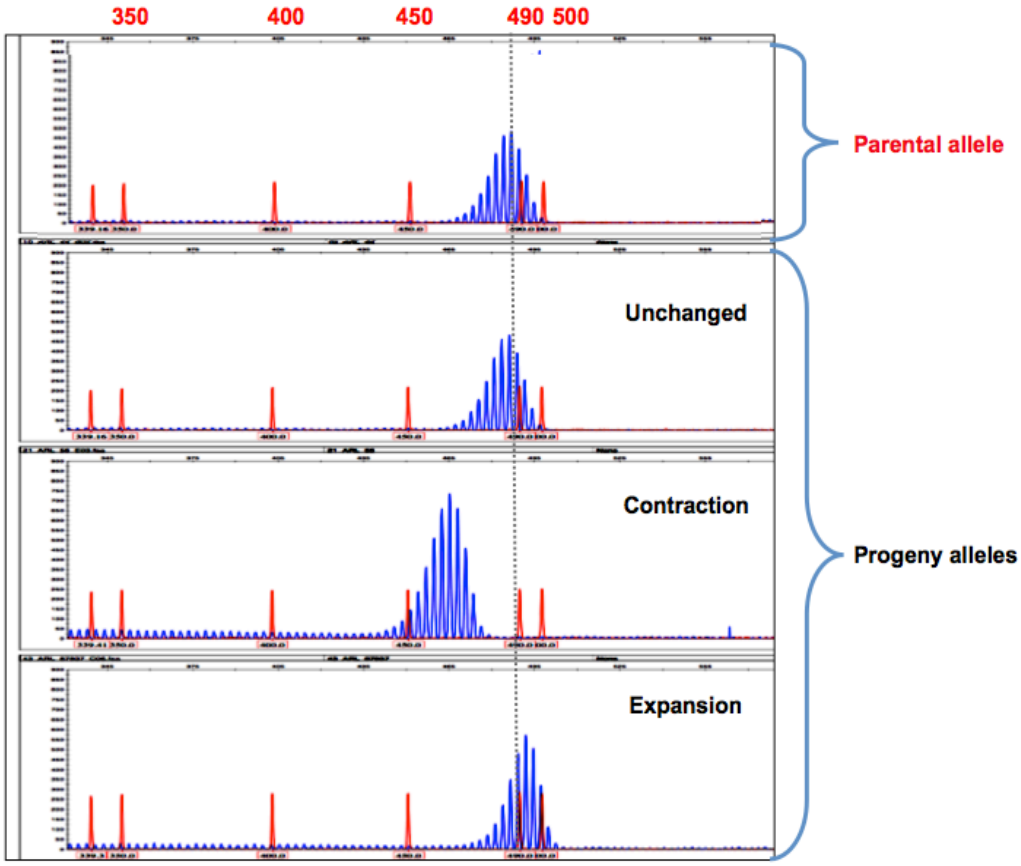


Figure 3.5 Example of GeneMapper profiles analysis. GeneMapper profiles illustrating intergenerational CGG•CCG repeat instability in FX PM mouse model .The first profile is the profile generated from the tail DNA of the parent, while the remaining profiles are profiles generated from the tail DNA of 3 of its offspring. The first pup has the same repeat number as the transmitting parent, thus the allele is scored as unchanged. The second pup has fewer repeats than the parent; therefore, it has undergone contraction. The last pup has more repeats than the parent. Thus this pup has inherited an expansion. Numbers above the GeneMapper profile corresponds to the sizes of different DNA fragments in the GeneScan™ 500 ROX™ molecular weight standard.

3.2.2.1 Intergenerational expansion is eliminated in MSH2 deficient mice

Analysis of offspring derived from *Msh2*$^{+/+}$ parents revealed that the frequency of expansions in the progeny was 96% when the PM allele was paternally transmitted and 62% when the allele was maternally transmitted (Figure 3.6). In the progeny of *Msh2*$^{+/-}$ mice the frequency of expansion was significantly reduced, to 72% for paternally transmitted alleles and to 38% when the allele was transmitted maternally.

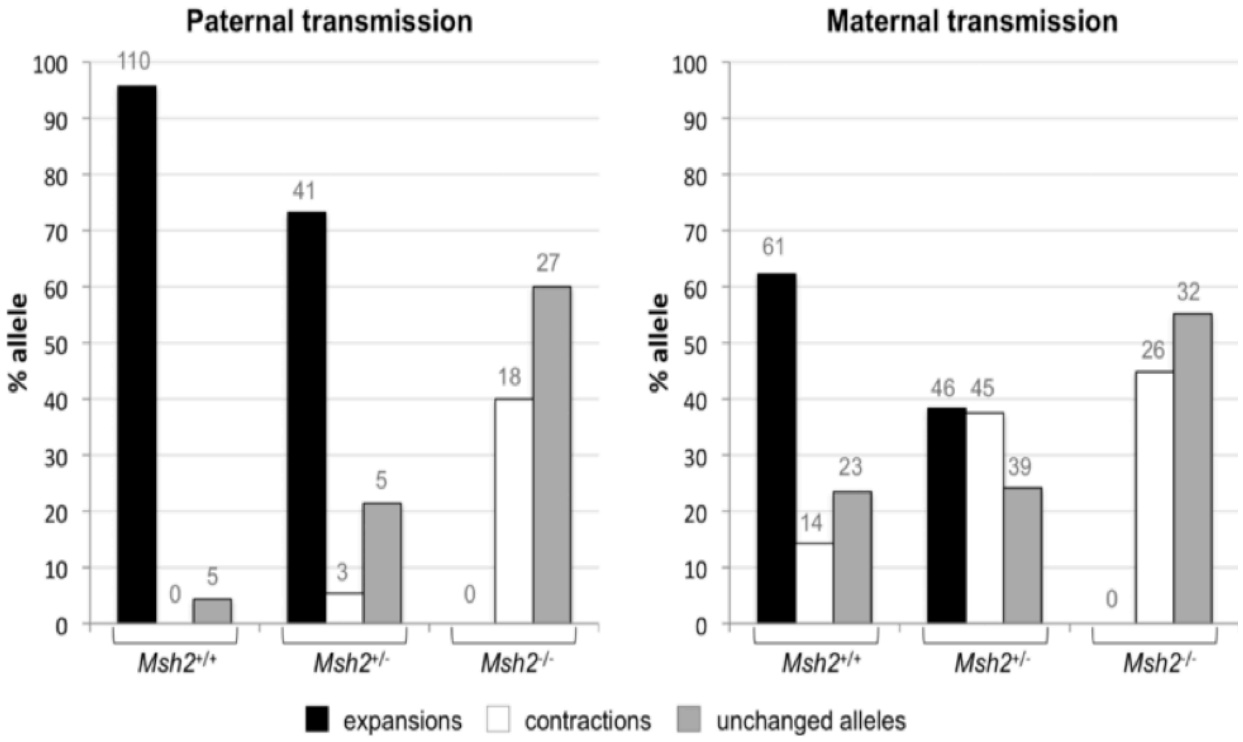


Figure 3.6 Effect of MSH2 deficiency on intergenerational instability. The frequency of expansions, contractions and unchanged alleles in the progeny of *Msh2*$^{+/+}$, *Msh2*$^{+/-}$,and *Msh2*$^{-/-}$ in FX PM mice. The numbers of animals analyzed in each category are shown above each bar.

The average size of the expansion in the progeny of *Msh2*$^{+/-}$ mice on paternal transmission was 5.5 repeats compared to 7.7 repeats in the progeny of *Msh2*$^{+/+}$ mice ($p<=0.009$) by t test. There was no statistically significant difference in the average

expansion size when the PM allele was maternally transmitted in *Msh2*$^{+/+}$ and *Msh2*$^{+/-}$ animals (3.2 compared to 3.0 repeats, p= 0.762 by t test (Figure 3.6).

The progeny of *Msh2*$^{-/-}$ mice showed no evidence of expansion on either maternal or paternal transmission of the PM allele. This loss of expansions was associated with an increase in alleles that had undergone contractions (40%) and alleles that were left unchanged (60%). These data suggest that MSH2 protein is essential and required for intergenerational repeat expansion of the PM alleles in the FX PM mice. Furthermore, the fact that in the progeny of *Msh2*$^{-/-}$ mice the number of contractions was increased, suggests that either that MSH2 is not necessary for intergenerational repeat contractions or that there is more than one contraction mechanism. Work by others in the Usdin laboratory suggest the latter, and that in the FX PM mouse model two contraction mechanisms exist one that requires MutSβ and one that is MSH2-independent (ZHAO *et al.* 2015).

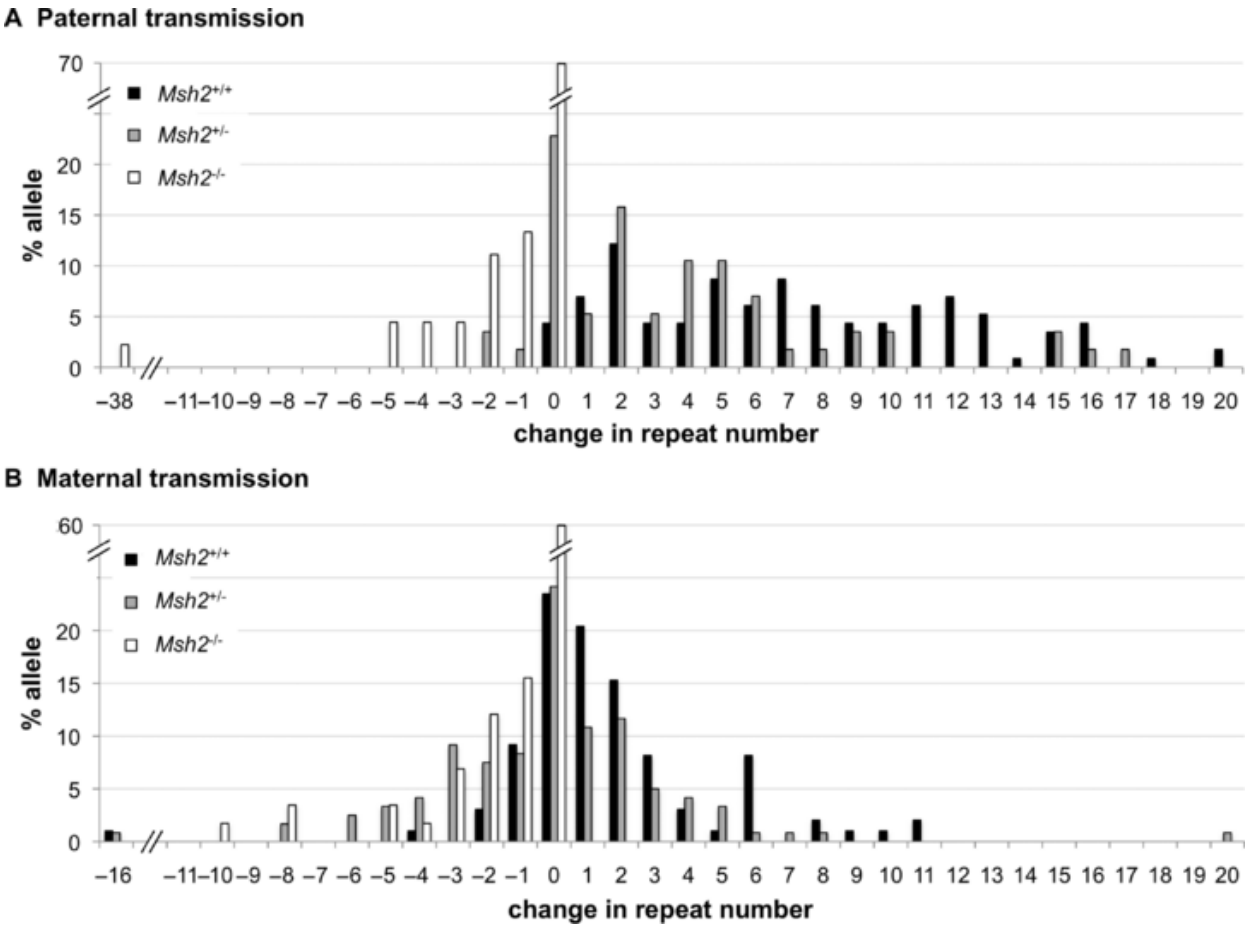


Figure 3.7 The profile of repeat size variations on intergenerational transmission in *Msh2*$^{+/+}$, *Msh2*$^{+/-}$, *Msh2*$^{-/-}$ mice. Graph showing the percentage of alleles with the indicated change in repeat length for each genotype for paternal (A) and maternal (B) transmission.

While no expansions were seen in the *Msh2*$^{-/-}$ progeny of *Msh2*$^{-/-}$ parents, *Msh2*$^{-/-}$ pups from *Msh2*$^{+/-}$ parents had a similar expansion frequency as their *Msh2*$^{+/+}$ littermates (Table 3.1). These data suggest that the expansion frequencies in the FX PM mice are dependent on the gene dosage of the transmitting parent instead of the offspring gene dosage.

Table 3.1 Frequency of expansion in progenies of *Msh2*$^{+/+}$, *Msh2*$^{+/-}$, and *Msh2*$^{-/-}$ parents

Transmitting parent	Parental genotypes*	Offspring genotype	# trans-missions	% expansions	% deletions	% un-changed
female	+/+	+/+	99	64	15	21
	+/-	+/+	25	48[§]	28	24
		+/-	56	43[§]	32	25
		-/-	19	37[§]	26	37
	-/-	**-/-**	**58**	**0**	**52**	**48**
male	+/+	+/+	115	96	0	4
	+/-	+/+	12	75[§]	8	17
		+/-	21	67[§]	5	28
		-/-	9	78[§]	0	22
	-/-	**-/-**	**45**	**0**	**40**	**60**

*Parents in each case had the same *Msh2* genotype, but only one had the PM allele. [§] Data not statistically significantly different

3.2.2.2 Intergenerational expansion is also reduced in MSH6 deficient mice

Having shown that in the absence of MSH2 all expansions are abolished in the FX PM mice; we then investigated whether MSH6, the partner of MSH2 in the MutSα complex involved in the repair of single-base mismatches, also played a role in intergenerational expansions. We found that the frequency of intergenerational expansions was significantly reduced in the progeny of *Msh6*$^{-/-}$ mice (Figure 3.8). On paternal transmission in *Msh6*$^{-/-}$ animals, we observed a lower frequency of expansion, 35% compared to 81% on transmission from *Msh6*$^{+/+}$ males ($p<0.001$). A higher frequency of contractions, 30% compared to 7% in *Msh6*$^{+/+}$ transmissions ($p<0.001$), was also seen. Similar results were seen for *Msh6*$^{-/-}$ maternal transmissions. The expansion frequency was 42% versus 85% in *Msh6*$^{+/+}$ animals ($p<000.1$) and the contraction frequency was 30% versus 8% in the progeny of *Msh6*$^{+/+}$ animals ($p<0.001$). We also observed an increase in the frequency of unchanged alleles in the *Msh6*$^{-/-}$ mice.

In contrast to what was seen in *Msh2* heterozygous mice, where the loss of one single functional *Msh2* allele significantly reduced the frequency of expansion, analysis of intergenerational instability in the progeny of *Msh6*$^{+/-}$ mice showed that the loss of one allele of *Msh6* was not sufficient to significantly reduce the frequency of expansion.

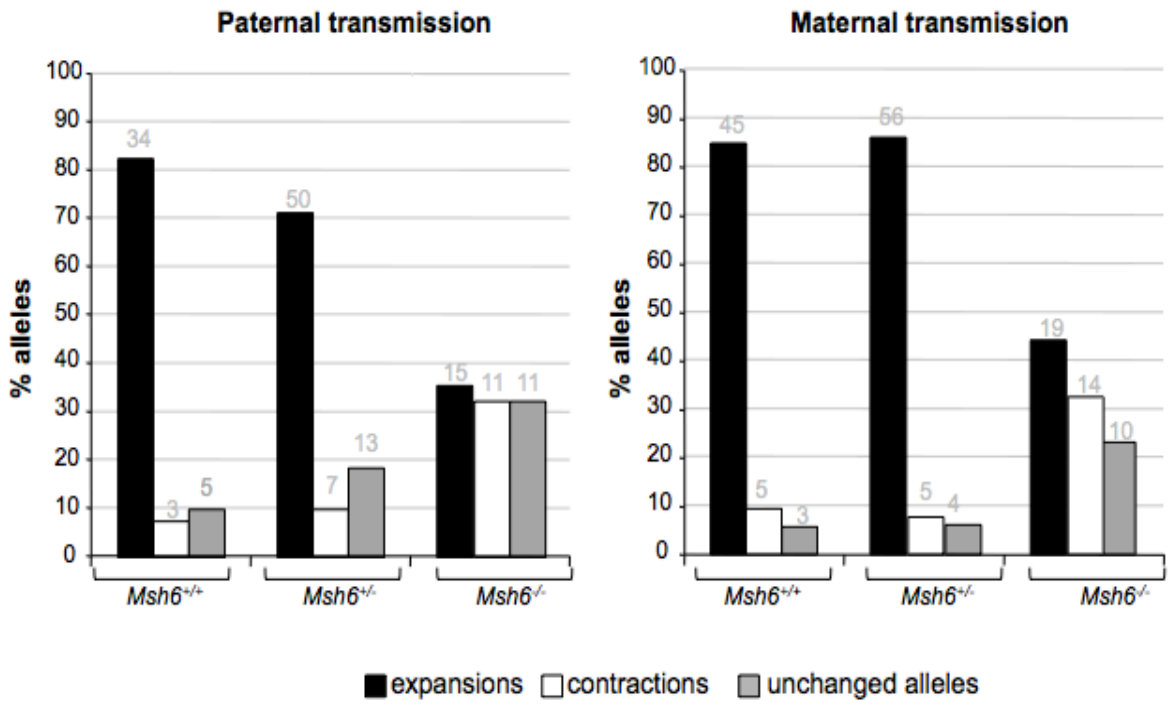


Figure 3.8 Effect of MSH6 deficiency on intergenerational instability. The frequency of expansions, contractions and unchanged alleles in the progeny of *Msh6*$^{+/+}$, *Msh6*$^{+/-}$,and *Msh6*$^{-/-}$ FX PM *mice.* The number above each bar represents the number of animals analyzed in each category.

The loss of MSH6 also significantly reduced the average number of repeats added (2.5 versus 5.7 repeats; $p<0.001$) on *Msh6*$^{-/-}$ paternal transmissions (Figure 3.9). In the progeny of *Msh6*$^{-/-}$ mothers the average number of repeats added was also reduced compared to *Msh6*$^{+/+}$ mothers, however, the difference was not statistically significant due to high standard deviation in WT samples (3.0 versus 5.5 repeats in *Msh6*$^{+/+}$ transmission; $p=0.096$). The repeat size was not reduced in the progeny of *Msh6*$^{+/-}$ mice on either paternal or maternal transmissions. Thus our data suggests that the loss of MSH6 reduces, but does not completely abolish, germ line expansions.

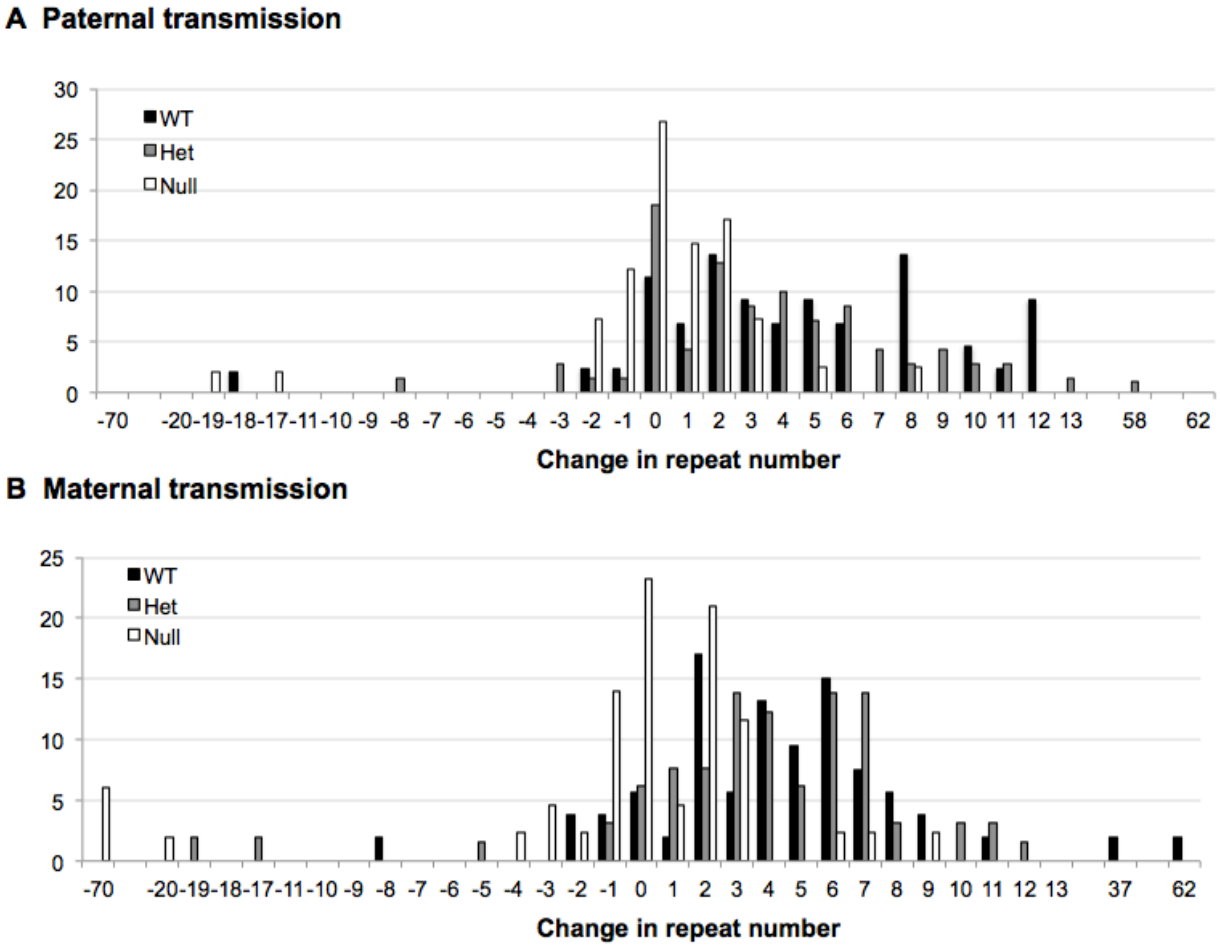


Figure 3.9 The profile of repeat size variations on intergenerational transmission in *Msh6*$^{+/+}$, *Msh6*$^{+/-}$, *Msh6*$^{-/-}$ mice. Graph showing the percentage of alleles with the indicated number of repeats added for each genotype for (A) paternal and (B) maternal transmissions.

To ascertain whether the reduced expansion frequency seen in different organs of *Msh6*$^{-/-}$ mice and on intergenerational transmission was an indirect effect of reduced levels of MSH3 in *Msh6* null mice as was suggested for in the DM1 transgenic mouse model on a mixed C57BL/6 x 129/OLA/FVB genetic background (FOIRY *et al.* 2006), we compared the levels of MSH2 and MSH3 in the ovary of *Msh6*$^{+/+}$ and *Msh6*$^{-/-}$ mice using Western blots. We found that, as expected, the level of MSH2 is reduced in the ovaries of *Msh6*$^{-/-}$ mice compared to *Msh6*$^{+/+}$ mice. This is consistent with the idea that the loss of MSH6 results in proteolytic degradation of the MSH2 normally bound to it. However, we did not find any significant differences in the level of MSH3 in *Msh6*$^{+/+}$ and *Msh6*$^{-/-}$ mice (Figure 3.10). Similar results for the levels of MSH3 were obtained in the brains of *Msh6*$^{+/+}$ and *Msh6*$^{-/-}$ of female mice using an ELISA assay (Table 3.2). However, the assay was not sensitive enough to use to verify the Western blot data for the MSH3 levels in ovary. Data generated by others in the laboratory showed that MSH3 is also not reduced in many other tissues of *Msh6*$^{-/-}$ male mice including brain, liver and testes, suggesting that the reduced expansion seen in the *Msh6*$^{-/-}$ FX PM mice is not due to an effect of reduced levels of MSH3 (thus the levels of MutSβ).

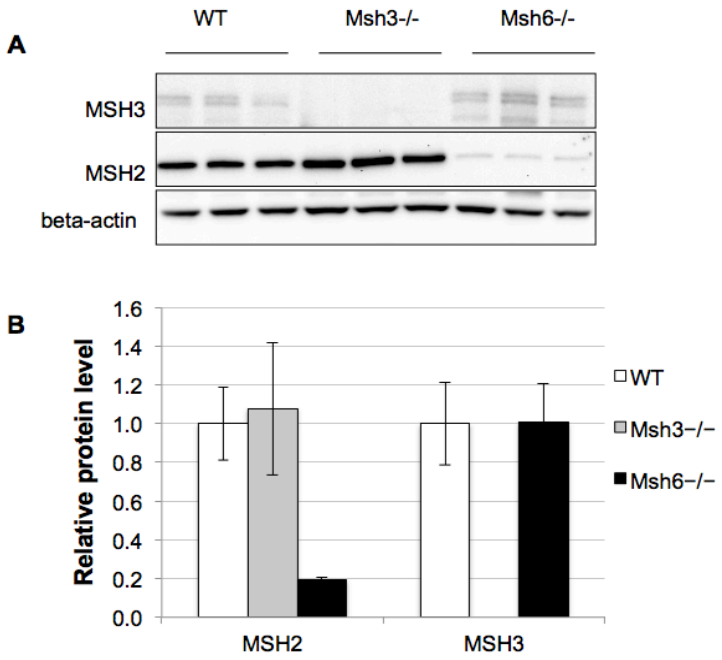


Figure 3.10 MSH2 and MSH3 protein levels in the ovaries of *Msh3* and *Msh6* null mice. (A) Western blot showing MSH2 and MSH3 levels in ovary (B) Relative amount of MSH2, and MSH3 in ovary of *Msh6*$^{+/+}$, *Msh3*$^{-/-}$ and *Msh6*$^{-/-}$ mice. Beta actin was used as protein loading control. In ovary, a doublet was observed for MSH3 that may indicate the presence of posttranslational modifications.

Table 3.2 MSH3 protein levels in the brain of *Msh6* null and *Msh6* WT mice

	Msh6$^{+/+}$			**Msh6$^{-/-}$**		
	Mouse 1	**Mouse 2**	**Mouse 3**	**Mouse 1**	**Mouse 2**	**Mouse 3**
Replicate 1	5.930	3.804	4.850	5.510	4.170	4.690
Replicate 2	5.920	3.550	4.990	5.620	4.260	4.880
Average (pg/ml)	5.925	3.677	4.920	5.565	4.215	4.7850
	Average (pg/ml)		SD	Average (pg/ml)		SD
	4.841		1.126	4.855		0.678

3.3 Discussion

To investigate the molecular mechanism involved in FX PM repeat instability in the FXDs, the effects of the MMR proteins MSH2 and MSH6 were examined in adult somatic tissues and during intergenerational transmission of the FX PM repeats. We report here that no evidence of expansion was seen at any age in any organ of *Msh2*$^{-/-}$ mice (Figure 3.1) or in any offspring of *Msh2*$^{-/-}$ mice regardless of the gender of the transmitting parent (Figure 3.6). Our data thus demonstrate that MSH2 is absolutely necessary for all CGG•CCG repeat expansions in the FX PM mice. Thus, our results differ from those seen in some other Repeat Expansion Disorders, where the loss of MSH2 had an effect on expansion frequency ranging from 0-96% depending on the sequence of the repeat and the gender of the transmitting parents (MANLEY *et al.* 1999b; KOVTUN AND MCMURRAY 2001; SAVOURET *et al.* 2003; WHEELER *et al.* 2003; DRAGILEVA *et al.* 2009; EZZATIZADEH *et al.* 2012). Furthermore, while all expanded alleles were replaced by unchanged and contracted alleles at similar frequencies in the FXDs, in other Repeat Expansion Disorders such DM1 where all expansions were replaced by contractions (~90%) (SAVOURET *et al.* 2003). Thus, our data reinforce the idea of how variable the effect of MMR components can be in different Repeat Expansion Disorders

We have also shown that the loss of a single functional *Msh2* allele significantly impacts both somatic and intergenerational expansions. The average expansion size in the most expansion prone tissues like brain and liver decreased significantly. However, in the testis, one of the other expansion prone tissues, the effect of loss of one *Msh2* allele was not statistically significant (Figure 3.2). We have shown in Chapter 2 that

MSH2 was more abundant in the testis than in the brain and liver (Figure 2.8). One explanation of our *Msh2*$^{+/-}$ data could be that in *Msh2*$^{+/-}$ mice, the MSH2 protein is rate limiting in brain and liver, but not in testis. Therefore, the difference in the repeat number added in *Msh2*$^{+/+}$ versus in *Msh2*$^{+/-}$ may suggest that the frequency at which expansion is occurring in the somatic cells is dependent on the levels of MSH2. As in somatic tissues, the loss of one functional *Msh2* allele resulted in a reduced frequency of expansion on both maternal and paternal transmissions (Figure 3.6). This is different from what is seen in the mouse model of CAG repeat expansion where loss of one *Msh2* allele did not reduce significantly the frequency of expansion in paternal transmissions (DRAGILEVA *et al.* 2009).

While no expansion was seen in any *Msh2*$^{-/-}$ progeny from *Msh2*$^{-/-}$ parents, a similar frequency of expansions was seen in the *Msh2*$^{-/-}$ progeny of *Msh2*$^{+/-}$ parents as their *Msh2*$^{+/+}$ or *Msh2*$^{+/-}$ littermates (Table 3.1). These data suggest that expansion frequencies may be sensitive to the gene dosage in the transmitting parent rather than in the offspring gene dosage. This would be consistent with previous evidence in the laboratory that expansions occur in the oocyte, a finding that is consistent with what has been proposed for humans (MALTER *et al.* 1997) and what can be inferred from the effect of maternal age on expansion risk in women (YRIGOLLEN *et al.* 2014)

Our data, thus, show that MSH2 is absolutely required for expansion in the FX PM mice. Furthermore, data obtained by others in the Usdin laboratory have shown that MSH3, the binding partner of MSH2 in MutSβ complex, is also required for all somatic expansions and more than 98% of germ line expansion in the FX PM mice (ZHAO *et al.*

2015). Taken together, our data suggest that the MutSβ complex is the major promoter of repeat expansions in the FXDs mice. However, my data shown here also demonstrates that MSH6 (therefore, MutSα) contributes significantly to both somatic and intergenerational expansions in the FX PM mice (Figures 3.3, 3.4, 3.8 and 3.9). I showed that this effect is not an indirect one resulting from changes in the levels of MSH3 protein in *Msh6*$^{-/-}$ mice (Figures 3.10, table 3.2). Moreover, while MutSα and MutSβ complexes do have some overlapping roles in the MMR pathway, it has been reported that loss of MSH6, thus, MutSα in mouse or in human cells does not affect the repair of MutSβ substrates (EDELMANN *et al.* 1997; OHZEKI *et al.* 1997; YANG *et al.* 2004; TSENG-ROGENSKI *et al.* 2012; CARETHERS *et al.* 2015). Thus, redeployment of MutSβ to damage sites normally repaired by MutSα is unlikely to account for the decrease in the intergenerational and somatic expansions seen in *Msh6*$^{-/-}$ mice. Thus, our data likely reflect a direct contribution of MutSα to the expansion process in the FX PM mice. This makes the FX PM mouse model the only Repeat Expansion Disease mouse model where MutSα has been found to promote both germ line and somatic expansions.

A role for MutSα in promoting repeat expansions is consistent with our previous observations in Chapter 2 that while MutSβ levels alone do not correlate well with the levels of somatic instability across the 5 different organs examined, a better correlation is seen when the levels of both MutSα and MutSβ are considered. However, since in the FX PM mouse almost all expansions are lost in the absence of MSH3 (ZHAO *et al.* 2015), MutSα must be somehow working in concert with MutSβ to generate repeat expansion in the FX PM mice. Work by others in the Usdin laboratory have

demonstrated that MutSα is able to stabilize the hairpins formed by the FX repeats that are thought to be the substrates for expansion (ZHAO *et al.* 2015). It may be that when MutSβ is rate limiting, that MutSα increases the half-life of the hairpins and prevents their removal by another repair mechanism, preserving them for later processing by MutSβ into expansions. It may be that the effect of MutSα is only apparent when the amount of the expansion substrate exceeds the ability of MutSβ to process them immediately.This could explain our observation that loss of MSH6 had a larger effect on somatic expansion in males than in females (Figures 3.3 and 3.4), since while MMR protein levels are very similar in males and females, males have higher expansion frequency (Chapter 2, Figure 2.4).

Since *Msh6*$^{-/-}$ animals (Figure 3.8), but not *Msh3*$^{-/-}$ animals (ZHAO *et al.* 2015), also show an increase in the contraction frequency, MutSα is also involved in protecting the genome against intergenerational contractions. Similar protective effects of MutSα against repeat contraction were reported in the mouse model of FRDA (EZZATIZADEH *et al.* 2012). The role of MutSα in protecting against contractions might be consistent with a classical MMR process, except that in this case the repair is triggered by an atypical repair substrate. Thus, our data suggest that MutSα participates in two different pathways, one that promotes expansion and another that protects the genome against contractions.

Chapter 4: Effect of a mutation in the BER gene, *PolB* on FX PM repeat expansion

4.1 Introduction

We have shown in the previous chapter that the MMR pathway is important for the expansion process in the FX PM mice. However, the mechanism by which the MMR pathway promotes repeat expansion remains unknown. Previous work in the Usdin laboratory that I was involved with prior to beginning my thesis, showed that treating FX PM mice with potassium bromate, a strong oxidizing agent, increases intergenerational repeat expansion (ENTEZAM *et al.* 2010). This raises the possibility that Base Excision Repair (BER), the major pathway involved in oxidative damage repair of DNA may also contribute to expansion. Since BER proteins are recruited to regions of open chromatin in response to oxidative stress (AMOUROUX *et al.* 2010), a role for BER in expansion might also explain why open chromatin is necessary for expansion as I demonstrated in Chapter 2. Previous work had demonstrated that the loss of OGG1 or NEIL1, two glycosylases that initiates BER, reduced somatic expansion in a mouse model of HD (KOVTUN *et al.* 2007; MOLLERSEN *et al.* 2012). However, loss of these proteins did not affect the frequency of germ line expansion. While it is possible that other DNA glycosylases may be more important for BER in the germ line, it was also possible that BER was not essential for expansion, but that nicks generated by BER DNA glycosylases, along with nicks generated by other enzymes, could be used by MMR proteins to subsequently generate expansions, perhaps by a form of non-canonical MMR repair (PENA-DIAZ *et al.* 2012).

In an effort to understand the mechanism of repeat expansion in the FXDs, we investigated the effect of an Y265C mutation in DNA polymerase beta (Polβ) on the

extent of expansion in somatic tissues and in sperm of FX PM mice. Polβ is a ~39kDa protein that belongs to X family of DNA polymerases that is essential for BER (LOEB AND MONNAT 2008) . Polβ possesses a 5'-deoxyribose-5-phosphate lyase activity that can remove native sugar phosphate residues generated by base excision. It is also able to carry out DNA synthesis to accomplish single nucleotide BER (BEARD AND WILSON 2006; LIU AND WILSON 2012). Polβ can also perform multi-nucleotide gap-filling and strand displacement synthesis that can result in an inefficient long-patch BER if the short patch is not efficiently completed due to oxidation or reduction of the deoxyribose 5-phosphate group (LIU AND WILSON 2012). It has been suggested that Polβ multi-nucleotide gap-filling synthesis can promote both expansion and deletion of the CAG repeat (LIU *et al.* 2009). Thus, this enzyme may play an essential role in modulating repeat instability.

The Y265C mutation is a dominant Polβ mutation found in humans that results in a hypomorphic Polβ variant with a slow DNA polymerase activity and a lower fidelity (SENEJANI *et al.* 2012). In order to investigate the role of Polβ deficiency on FX repeat expansion, we examined somatic and germ line instability in the sperm of mice in which the normal mouse *PolB* gene was replaced with one carrying the Y265C mutation (SENEJANI *et al.*).

4.2 Results:

It had been reported that approximately 40% of *PolB* homozygous mutant mice (*PolB*$^{c/c}$) survive past birth. However, in our hands we were unable to generate any viable homozygous mutant mice in a pure C57BL/6 background and in animals backcrossed for 4 generations onto a 129S1 background. Consequently, we limited our research to heterozygous (*PolB*$^{+/c}$) animals only. We assessed somatic instability in mice at 16 months of age and germ line instability in the sperm of 3 and 11 month old animals.

4.2.1 Somatic instability is reduced in mice with a Y265C mutation in Polβ.

In order examine the role of Polβ Y265C mutation on somatic expansion, we first generated mice that carried the PM allele and that were heterozygous for the Y265C mutation by crossing our FX PM mice to mice with Y265C mutant mice. We then evaluated the degree of expansion in different tissues of 16-month-old *PolB*$^{+/+}$ and *PolB*$^{+/C}$ males with ~140 repeats in their *Fmr1* gene. The genomic DNA from tail, brain, kidney, liver, heart, testes was amplified by fluorescent PCR, resolved on a 3130XL Genetic Analyzer and analyzed using GeneMapper 4.0 software, as detailed in the Materials and Methods chapter. Somatic expansion was analyzed by comparing the repeat PCR profile of each organ to the repeat profile seen in the tail DNA of the same animal at three weeks of age (tail 1) and by quantifying the amount of somatic expansion using the somatic instability index (SII), a quantitative analysis of the degree of somatic instability developed in the Wheeler laboratory (LEE *et al.* 2010).

Analysis of the data from three 16 month-old-male *PolB*$^{+/+}$ and three 16 month old *PolB*$^{+/C}$ mice revealed that heterozygosity for the *PolB* Y265C mutation significantly reduced the extent of somatic expansion in some expansion prone tissues (Figure 4.1). While the repeat was stable in the heart of both *PolB*$^{+/+}$ and *PolB*$^{+/C}$ mice, consistent with our previous findings described in chapters 2 and 3, the extent of expansion was significantly reduced in the tail (p=0.013) and in the testes (p=0.0011) of *PolB*$^{+/C}$ mice. We also observed a slight decrease in the extent of somatic expansion in the kidney. However, this decrease was not statistically significant (p= 0.07). No difference was seen between SII values in the liver or brain of *PolB*$^{+/+}$ and *PolB*$^{+/C}$ mice. Since we have shown in chapters 2 and 3 that somatic contractions do not occur in the mouse model of FX PM, the reduced SII seen in the tail and testes of *PolB* Y265C mutant mice is consistent with a role of Polβ in promoting somatic expansions in our mouse model. The fact that heterozygosity for the *PolB* mutation affects expansion in only a few tissues would be consistent with the observation that loss of one copy of *PolB* does not affect mutation rates in most tissues, maybe because some cells have regulatory mechanisms that can compensate for the presence of only one functional allele (ALLEN *et al.* 2008).

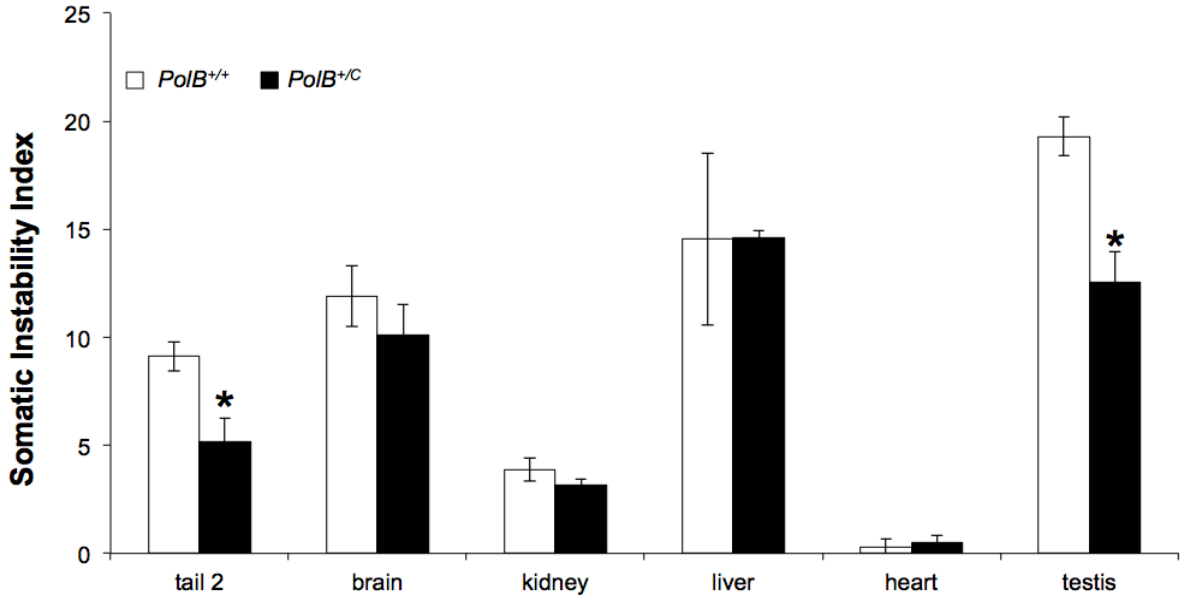


Figure 4.1 Expansion as assessed by the SII in indicated organs of *PolB*$^{+/+}$ and *PolB*$^{+/C}$ in male FX PM mice. The average somatic instability index (LEE *et al.* 2010) for the indicated organs. WT (black bars) n= 3 *PolB*$^{+/+}$ and *PolB*$^{+/C}$ n= 3 (white bars) at 16 months of age. The asterisks indicate organs in which the Y265C mutations have a p value< 0.01**).**

4.2.2 Germline expansion is also reduced in mice with a Y265C mutation in Polβ

In order to examine the role of *PolB* Y265C mutation on germ line instability, we used small pool PCR of sperm DNA in 3 and 11 month old heterozygous males carrying the PM. We isolated the genomic DNA from sperm as described in the Materials and Methods chapter and carried out small pool PCR on these DNAs using a nested PCR strategy. The first round of PCR was carried out with ~ 3pg of genomic DNA as template, the equivalent of one haploid mouse genome, using primer pair Frax C and Frax F. No PCR product was visible on agarose gel electrophoresis at this stage. The second round of PCR was done using 1 µl of the DNA generated from the first round PCR and primer pair Frax m5 and Frax m4 (Frax m4- labeled with 6-carboxyfluorescein

(FAM)). The PCR product from the second round was then resolved on an ABI 3130XL Genetic Analyzer and the data analyzed using GeneMapper 4.0 software, as detailed in the Materials and Methods chapter.

Analysis of 3-month-old animals showed that the number of expanded alleles observed was significantly lower in the sperm of *PolB*$^{+/C}$ mice than in the sperm of *PolB*$^{+/+}$ mice. Only 42% of the alleles from gametes of *PolB*$^{+/C}$ mice were larger than the parental allele compared to 74% in *PolB*$^{+/+}$ mice (p=0.0001) (Figure 4.2). We also observed an increase in proportion of contractions in the gametes of *PolB*$^{+/C}$ mice (37% versus 7% of alleles, p=0.0001). The proportion of unchanged alleles, alleles that were the same size as the parental allele, was similar in both *PolB*$^{+/C}$ and *PolB*$^{+/+}$ mice (21% versus 18 %).

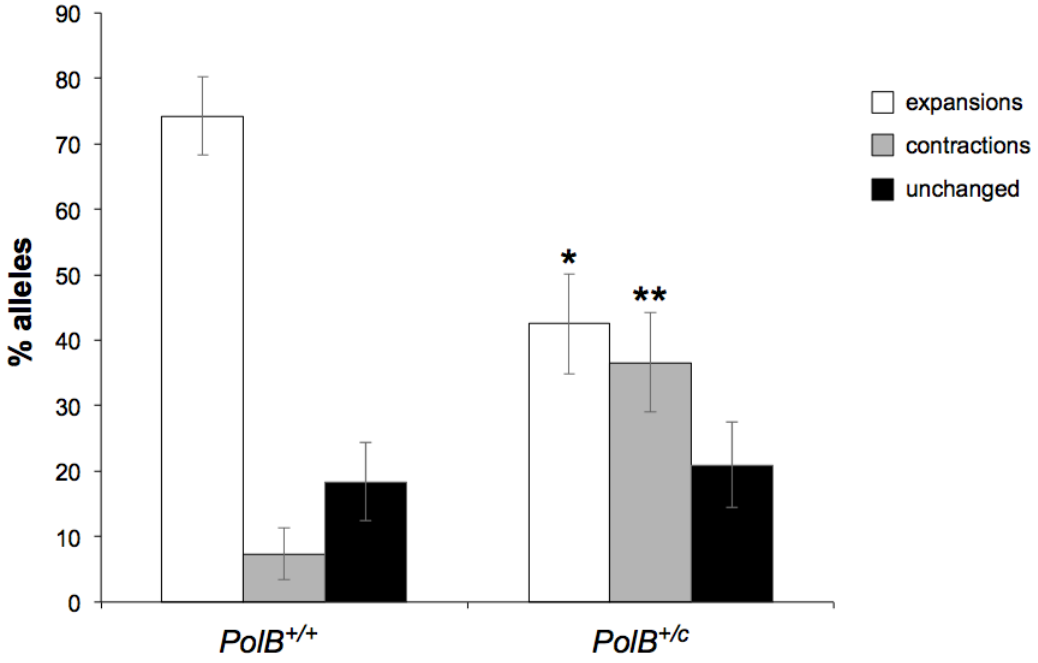


Figure 4.2 Role of Y265C mutation in *PolB* on intergenerational instability in gametes of 3 months mice. The proportion of expansions, contractions and unchanged alleles in *PolB*$^{+/+}$ and *PolB*$^{+/C}$ male mice. The 95% confidence interval is represented by the error bars. The asterisks indicate the P value by Fisher's exact test (*p <0.01 and ** p<0.0001).

While the total number of expansions was reduced, we found that *PolB*$^{+/C}$ mice had more larger expansions with 22 % of expansions involving the addition of more than 10 repeats in the gametes of *PolB*$^{+/C}$ mice compared to only 3% in the gametes of wild type mice (Figure 4.3). We also noticed that the size of contractions was larger in *PolB*$^{+/C}$ mice compared to wild type mice (Figure 4.3). The average size of contractions was 19.68 repeats in the sperm of *PolB*$^{+/C}$ mice and 10.83 repeats in the sperm of *PolB*$^{+/+}$ mice. However, the difference between the two genotypes was not statistically

significant because of some very large contractions that were seen in the *PolB*$^{+/C}$ mice that resulted in a very high standard deviation (SD=31.58 versus 10.05 *PolB*$^{+/+}$).

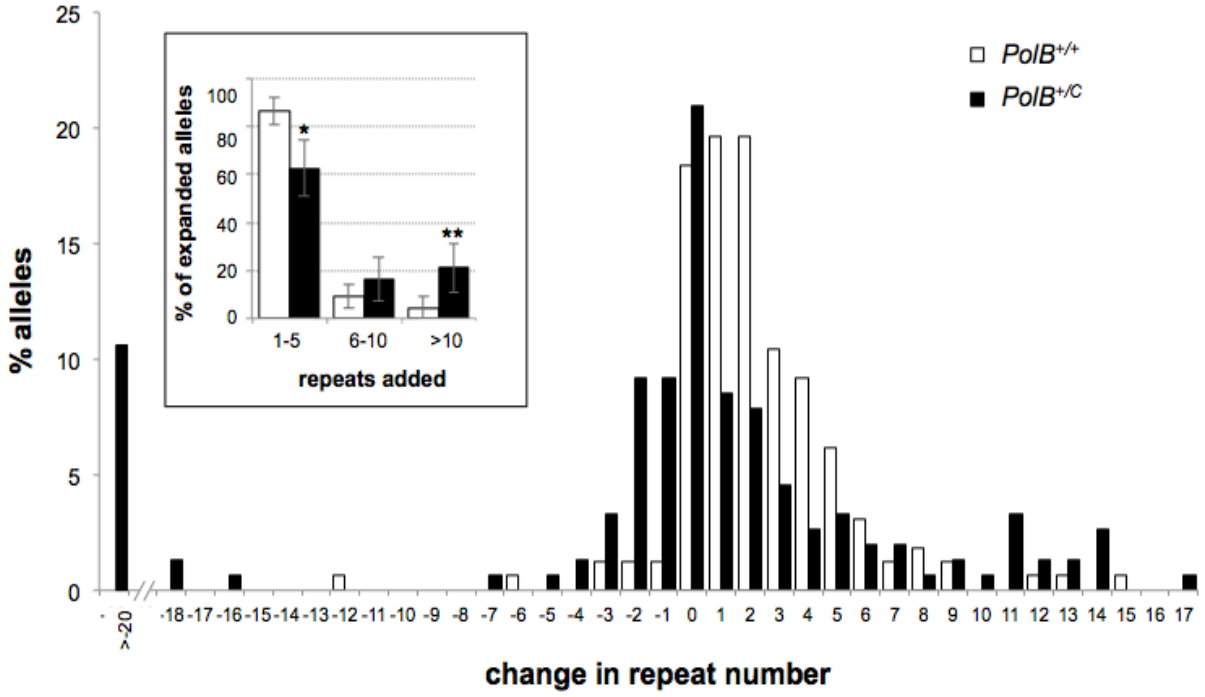


Figure 4.3 The profile of repeat length changes in the gametes of *PolB*$^{+/+}$ and *PolB*$^{+/C}$ young mice. Graph showing the proportion of alleles with the indicated change in repeat length for *PolB*$^{+/+}$ and *PolB*$^{+/C}$ gametes of 3-month-old male mice. Inset: Graph showing different classes of repeat size in *PolB*$^{+/+}$ and *PolB*$^{+/C}$ mice. The 95% confidence interval is represented by the error bars. The asterisks indicate the allele classes that are significantly different by Fisher's exact test (*p < 0.01 and ** p< 0.001)

In the sperm of 11-month-old *PolB*$^{+/+}$ and *PolB*$^{+/C}$ mice more than 90% of alleles had expanded (Figure 4.4). Since in 3 month old *PolB*$^{+/C}$ mice, the proportion of gametes that had undergone at least one round of expansion was 42%, the fact that in 11 month old *PolB*$^{+/C}$ mice ~90% of gametes revealed evidence of at least one round of expansion was not unexpected.

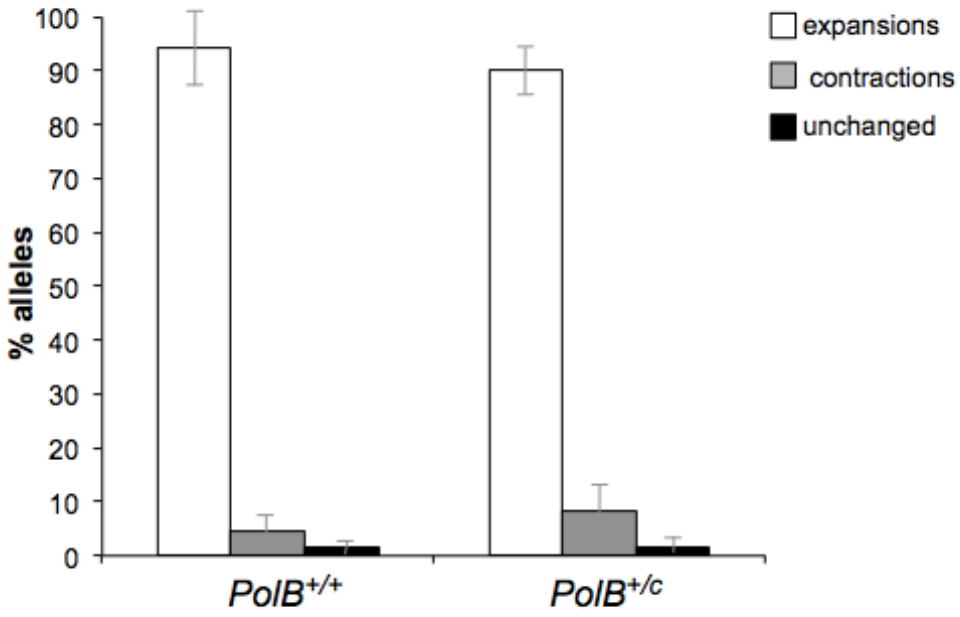


Figure 4.4 Role of Y265C mutation in *PolB* on intergenerational instability in the gametes of 11 months mice. The proportion of expansions, contractions and unchanged alleles in *PolB*$^{+/+}$ and *PolB*$^{+/C}$ male mice. The 95% confidence interval is represented by the error bars.

The excess of large expansions seen in 3 month old *PolB*$^{+/C}$ mice was even greater in the gametes of 11 month old animals (Figure 4.5) with 39% of the expanded gametes having added >15 repeats compared to 9 % in *PolB*$^{+/+}$ mice (Figure 4.5, inset). Furthermore, in the gametes of *PolB*$^{+/C}$ mice, the distribution profile of all expansions showed local maxima representing the gain of 10,16, 21, and 27 repeats (Figure 4.5). This pattern would be in agreement with the idea that most expansions in the gametes of *PolB*$^{+/C}$ mice result in the addition of an average of 5 to 6 repeats, with the gametes of 11-month-old animals already having undergone more than one round of expansion. In contrast, the average number of repeats added in *PolB*$^{+/+}$ mice was 1-2 repeats. Our data suggest that while the Y256C mutation lowers the frequency of repeat expansion, the residual expansions are larger.

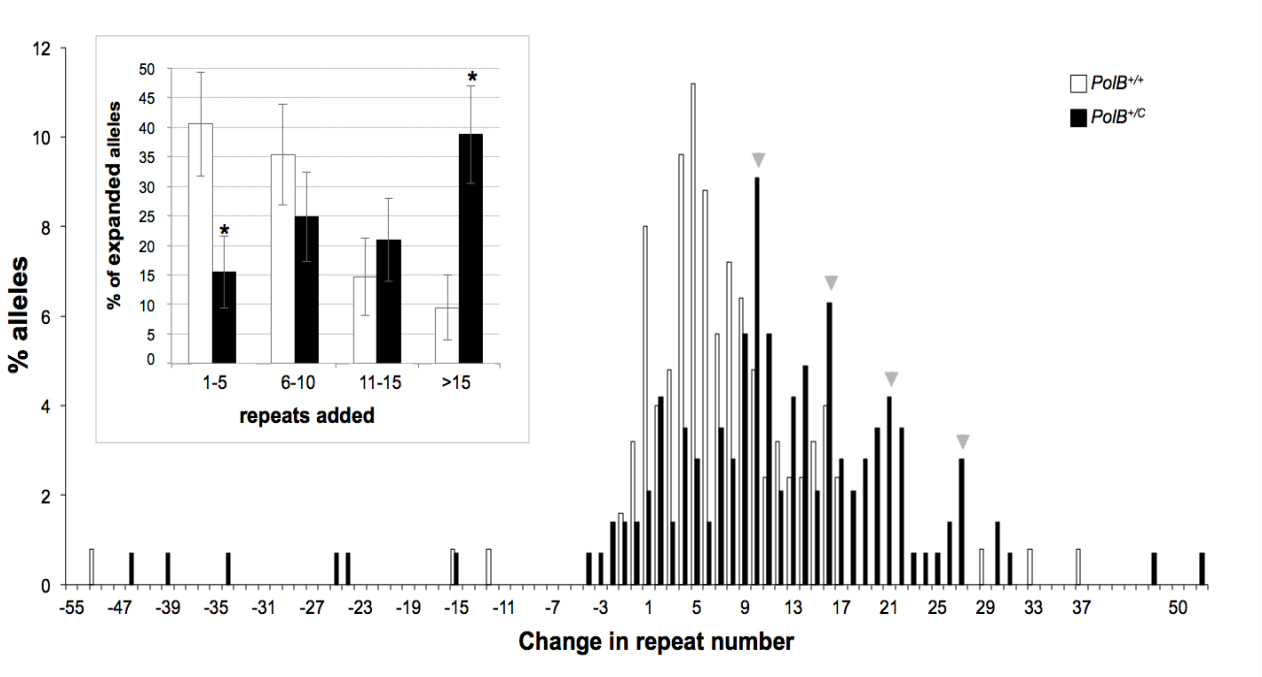


Figure 4.5 The profile of repeat length changes in germ line of gametes of *PolB*$^{+/+}$ and *PolB*$^{+/C}$ older mice. Graph showing the proportion of alleles with the indicated change in repeat length for *PolB*$^{+/+}$ and *PolB*$^{+/C}$ gametes of 11-month-old male mice. Inset: Graph showing different classes of repeat size in *PolB*$^{+/+}$ and *PolB*$^{+/C}$ mice. The 95% confidence interval is represented by the error bars. The asterisks indicate the allele classes that are significantly different by Fisher's exact test (*$p < 0.01$). Arrowheads represent the local maxima as observed in the distribution of *PolB*$^{+/C}$ alleles.

4.3 Discussion:

The aim of this research was to assess whether events downstream of DNA nicking in the BER pathway are involved in the expansion of the FX PM repeats. We have shown that heterozygous mutation for the *PolB* Y265C mutation causes a significant decrease in the extent of somatic expansion in 16 month old mice (Figure 4.1). However, not all expansion-prone tissues tested were affected by the presence of the *PolB* Y265C mutation, expansions were reduced in the tail taken at euthanasia and the testis only. This is consistent with previous findings on the limited effect of *PolB* mutations on the mutation rates of different tissues (ALLEN *et al.* 2008), an observation that has been interpreted to mean that DNA repair in most tissues is not sensitive to *PolB* heterozygosity. We also have shown that heterozygosity for the *PolB* Y265C mutation significantly reduces the intergenerational expansion frequency in young FX PM mice (Figures 4.2 and 4.3). While the number of expansions seen in older *PolB* heterozygous mice was not significantly different from WT animals (Figure 4.4), this is likely due to the fact that the number of expansions has reached a maximum.

While the mechanism by which BER pathway promotes repeat expansion remains to be elucidated, it has been suggested based on in vitro experiments that expansion can arise during long patch BER either via a “hit and run” mechanism involving limited Polβ strand displacement and FEN1 flap excision of the nucleotide linked to the 5′-sugar phosphate group as illustrated in Figure 4.6i. The use of such a pathway might be expected to produce only small expansions since Polβ is not very processive and does not result in extensive strand displacement synthesis. An alternate

model suggests that expansion result from the use of an alternate LP-BER sub-pathway that involves the coordinated DNA synthesis by Polβ and Polδ or Polε, followed by FEN1 flap cleavage (LIU AND WILSON 2012; CHAN *et al.* 2013) as illustrated in Figure 4.6ii. Since Polδ is more processive it is more likely to result in strand-slippage. In the presence of Polβ priming may be more likely to occur from the slipped position since Polβ lacks the appropriate proofreading ability necessary for removal of the slipped strand. In addition, Polδ causes more extensive strand displacement than Polβ. Thus, the use of this pathway may be expected to produce larger expansions. Our data on the differential effect of the Y265C mutation on the occurrence of large and small expansions (Figures 4.3 and 4.5) might reflect the contribution of these different mechanisms to repeat expansion in the FX PM mice. It is possible that the small expansions (Figure 4.3) observed in our mice result from the incorporation of few additional nucleotide bases into DNA strand during repair by the LP-BER branch that is dependent on the weak strand-displacement activity of Polβ as illustrated in Figure 4.6i, whereas large expansions may arise from the incorporations of several nucleotide bases in the repaired strand by LP-BER branch that is dependent on more extensive strand displacement by Polδ/Polε as shown in Figure 4.6ii. Our data suggest that expansion in the FX PM mouse involves late events in the BER pathway.

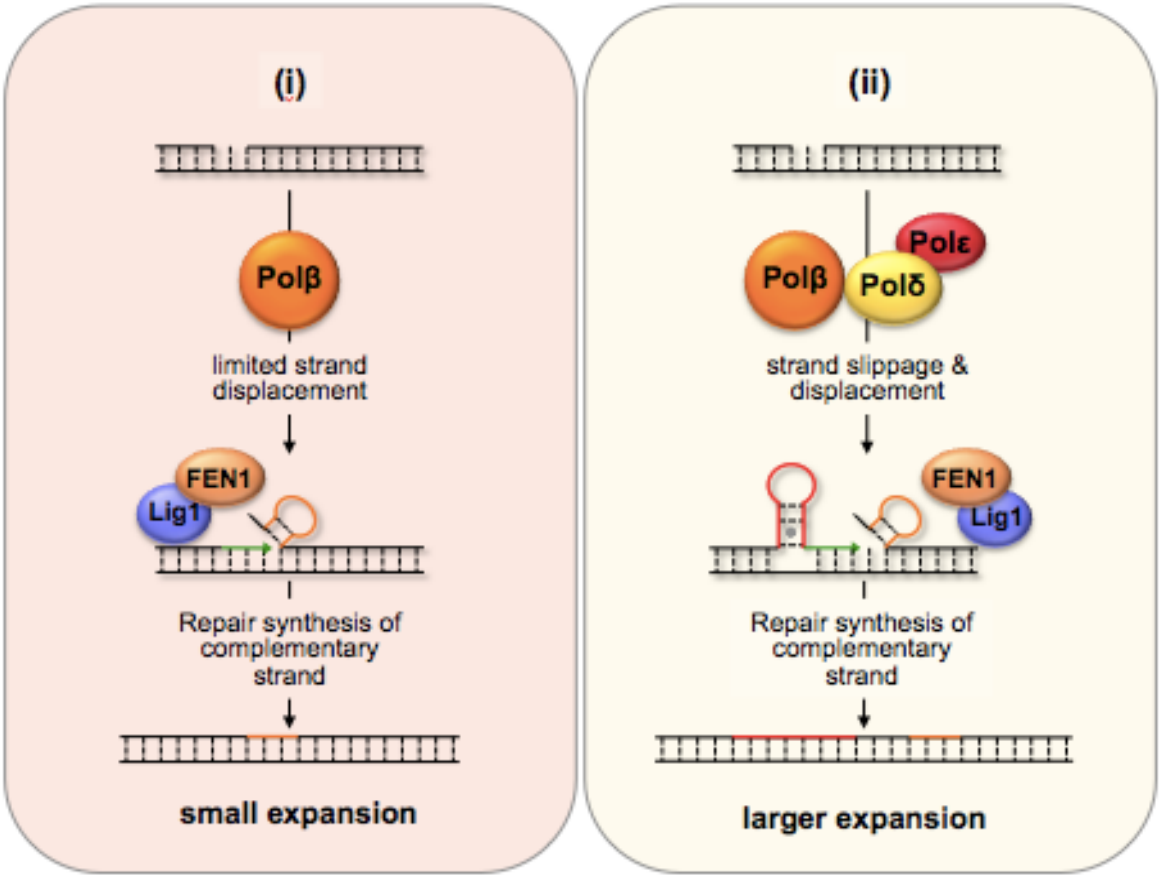


Figure 4.6 BER-model of expansion in the PM mouse of the FXDs. The repair of DNA damage within the repeat initiated by DNA glycosylases in response to oxidized bases. The short patch BER sub-pathway can repair a native 5' -sugar phosphate group. Nicks unrepaired by SP- BER will be processed by either one of the two sub-pathways of the LP BER: (i) In one LP BER pathway (left), the processing of nicks is done by Polβ gap-filling synthesis, followed by cleavage of the small DNA flap by FEN1. A 5'end that can be ligated is generated and still contains few additional bases. This pathway is also referred to as "hit and run" mechanism of BER and generates small expansions. (ii) The LP BER pathway (right) that involves the use of Polβ or Polδ, and Polε. In this pathway repair is mediated by an inefficient strand displacement by Polβ or more extensive strand displacement by Polδ or Polε, followed by the creation of a ligatable 5' end that also still contains more flap bases by FEN1 (shown in orange). This pathway generates larger expansions. Modified from Zhao and Usdin, 2015.

Chapter 5: Conclusions

5.1 Conclusions

The aim of my thesis work was to more fully characterize the somatic instability of PM alleles and examine the involvement of DNA repair proteins MSH2, MSH6, and Polβ in CGG•CCG repeat expansion in somatic cells and on intergenerational transmissions in the FX PM mouse model. In this thesis, I described factors important for CGG•CCG repeat expansion using this mouse model and a lymphoblastoid cell line from a human PM carrier. The novel discoveries of this research thesis have been categorized into four sections:

Firstly, we showed for the first time that somatic expansion could occur in human PM cells in tissue culture converting the PM allele gradually over a period of time of ~2 years into a FM allele (Figure 2.2). Somatic expansion may have clinical significance since the gravity of the disease symptoms is related to the number of repeats.

Secondly, we showed that somatic expansion in the FX PM mouse was more prominent in male than in female mice and that this gender difference was due in part to X chromosome inactivation in females (Figure 2.9). Specifically, we showed that somatic expansion occurred exclusively on PM allele that was located on the active X chromosome of female mice, suggesting that either an open chromatin configuration or the transcriptional activity of the gene is necessary for expansion. In fact, transcription has been previously proposed to play a role in repeat expansion based on observations made of the expansion rates seen in different lines of transgenic mouse models of DM1 and HD (LIA *et al.* 1998; GOULA *et al.* 2012). However, since chromosomal context

affects expansion and the chromosomal context of the transgenes being compared in these cases differed from one another, it was not possible to definitively implicate transcription/open chromatin in expansion based on these experiments. However, since *Fmr1* is an X-linked gene, we were able to compare simultaneously the extent of expansion of the same PM allele in the same sequence context on the active and inactive X chromosome in the same animal. Thus our data allowed us to definitively demonstrate a requirement for transcription and/or an open chromatin configuration in the FX expansion in mice. A role for transcriptionally competent chromatin is also consistent with the observation that large, unmethylated human alleles are unstable (GLASER *et al.* 1999) and suggests that at least some of this instability reflects expansions of smaller alleles rather than contraction of larger ones.

However, when colleagues in the Usdin laboratory measured the amounts of *Fmr1* mRNA in five different expansion prone tissues, no good correlation was seen between the amount of transcription and the degree of somatic expansion in our mice (LOKANGA *et al.* 2013). It might be that while transcription or an open chromatin conformation is essential, it is not enough by itself to promote expansion; other factors or processes associated with transcription might be rate-limiting.

Thirdly, my work showing that expansion of the CGG•CCG repeat in our mice was dependent on the parental *Msh2* gene dosage and not that of the offspring. This suggests that expansion of the FX repeat most likely occurred in the gametes. Given that the oocyte does not divide, this lends support to the idea that the expansion mechanism does not involve a problem with chromosome replication per se and that a

mechanism involving aberrant DNA repair and/or recombination is likely to account for repeat expansion. I also demonstrated the involvement of MutSα in the generation of both germ line and somatic expansion in the FX PM mice. However, since data generated by colleagues in the Usdin laboratory have shown that loss of MSH3 (therefore, MutSβ) abolished almost all expansions (ZHAO *et al.* 2015), MutSα must be acting in a predominantly MutSβ-dependent manner.

Fourthly, I also showed that expansions were significantly reduced in the paternal gametes of young mice (Figure 4.2) and in the tail and testis tissues of FXD mice that were heterozygous for a Y265C mutation in *PolB* gene (Figure 4.1). Thus expansion involves central events in the BER pathway. We have shown in chapter 3 of this thesis that MSH2 and MSH6 (thus, MutSα) contribute to the generation of expansions in our mouse model. Thus, it might be that these two pathways interact to generate expansion. However, the mode of interaction between the components of these two pathways to generate expansions remains an open question. It might be that MutSα is acting downstream of DNA damage excision to stabilize the hairpins formed during strand-slippage and strand-displacement synthesis that is mediated by Polβ or Polδ, and Polε. It can also perhaps interfere with the removal of these hairpins by enzymes like FEN1 (SPIRO *et al.* 1999). This interference could result in the hairpins being channeled into a different repair pathway that results in expansions (ZHAO AND USDIN 2015)

Thus our findings lend credence to the idea that MMR and BER pathways interact to promote expansion (KOVTUN *et al.* 2007; MCMURRAY 2010). However, more work is needed to understand just how these pathways interact to generate expansions

in the FXDs in particular, and in the Repeat Expansion Disorders in general. Taken all together, my research findings have shed significant new light on the mechanisms of repeat expansion in the FXDs.

Chapter 6: Materials and Methods

6.1 Mouse maintenance and breeding

Generation of the FXPM mice and of *Msh2*, *Msh6*, and *PolB* mutant mice was described previously (REITMAIR *et al.* 1995; EDELMANN *et al.* 1997; ENTEZAM *et al.* 2007; SENEJANI *et al.* 2012). Mice were kept on a predominantly C57BL/6J genetic background. However, in the effort to generate *PolB* homozygous mutant mice, we also backcrossed some mice for 4 generations onto a 129S1 background. They were all maintained in accordance with "the guidelines of the NIDDK Animal Care and Use Committee and with the *Guide for the Care and Use of Laboratory Animals"* (NIH publication no. 85-23, revised 1996). Newborn mice were weaned at 3 weeks of age and mice were tagged with a number for identification on the ear. Tail snips (~<5 mm) were taken at this time using local anesthesia with ethyl chloride. Mice were euthanized when necessary by CO_2 inhalation followed by cervical dislocation.

To generate FXPM mice with mutations in either MMR gene (*Msh2* and *Msh6*) or the BER gene, *POLB*, we crossed FX PM mice with ~147 repeats with mice carrying null alleles (knockouts) of *Msh2* and *Msh6* and mice having a hypomorphic mutation in *POLB* to generate animals carrying a single copy of the *Fmr1* PM allele that were wild type (WT), heterozygous, or homozygous for *Msh2* or *Msh6* and heterozygous for the *POLB* mutation.

6.2 DNA isolation

Genomic DNA was extracted from mouse-tail samples using KAPA genotyping kits (Cat # AS1120; KAPA BIOSYSTEMS, Wilmington, MA) according to the manufacturer's recommendations. The salting out method for isolation DNA from tail and sperms was also sometimes used. Briefly, 180 µl of tissue lysis buffer (buffer ATL # 19076, Qiagen, Valencia, CA) and 20 µl of proteinase K (50 mg/ml) were added to the tails in 1.7 ml microfuge tubes. The tubes were briefly vortexed and incubated overnight at 55°C. The next day, the tubes were briefly centrifuged to bring any liquid to the bottom of tubes and 60 µl of 5 M NaCl was then added. Subsequently, samples were mixed well by inverting the tubes 10 times, and then spun down at maximum speed (~14000 rpm) for 10 minutes at ambient temperature (~73°F). Afterward, 260 µl of the supernatant was transferred to a new microfuge tube and 260 µl of 100% ethanol was added to the supernatant. The samples were once more spun down for 10 minutes. After the removal of the ethanol, the DNA pellets were then washed with 1 ml of 70% ethanol. The DNA pellets were then air dried for about 10 minutes before being resuspended in 50-100 µl of TE buffer. DNA samples were stored at 4°C. A DNA isolation kit (Maxwell 16, Promega, Madison, WI) was also used for the extraction of genomic DNA from different mouse tissues.

The concentrations of the genomic DNA were determined using a Denovix DS-11 spectrophotometer (Denovix, Wilmington, DE). The concentration of samples was determined based on their absorbance (A) at 260 nm and their purity was verified by analyzing A260/280 ratio. In order to verify the integrity of the genomic DNA, we

electrophoresed ~100 ng of the genomic DNA on a 0.7% agarose gel along with λ DNA/Hind III molecular weight marker (cat # 15612-013; Life Technologies, Grand Island, NY). After staining with ethidium bromide the DNA was visualized on a FluorChem M gel imager (Proteinsimple, San Jose, CA).

6.3 Genotyping

The primers used for genotyping are listed in Table 1. Primers were all obtained from Integrated DNA Technologies (IDT, Coralville, IA). The PCR conditions are summarized in Table 2. FX PM genotyping in mice were carried out using a PCR assay as previously described (LAVEDAN *et al.* 1998) and shown in Table 2. The binding sites for the primers used in this assay are found immediately adjacent to the repeat tract and their 3' ends are specifically unique to the PM allele. Another primer pair, Frax C and Frax F, that detect both wild type (WT) *Fmr1* and FX premutation alleles, was used for genotyping (FU *et al.* 1991). For the analysis of repeat size in the human cell line, we used Frax AR and Frax F primer pair. Frax F was labeled with 4, 7, 2′, 4′, 5′, 7′-hexachloro-6-carboxyfluorescein (HEX). The PCR products were separated by electrophoresis on a 1.5% agarose gel, stained and visualized as described above.

Msh2, *Msh6*, and *POLB* genotyping was also carried out by Polymerase Chain Reaction (PCR) analysis of tail genomic DNA, as described previously (REITMAIR *et al.* 1995; EDELMANN *et al.* 1997; SENEJANI *et al.* 2012) and shown in Table 2. The PCR products were resolved by electrophoresis on a 1.5% agarose gel, stained and visualized as described above.

Table 6.1 Primers used for genotyping.

PCR	Primer ID	Sequence (5'-3')	References
CGG	Frax-m4*	CTTGAGGCCCAGCCGCCGTCGGCC	(Entezam, et al., 2007)
	Frax-m5	CGGGGGGCGTGCGGTAACGGCCCAA	
	Frax C	GCTCAGCTCCGTTTCGGTTTCACTTCCGGT	
	Frax-F*	AGCCCCGCACTTCCACCACCAGTCTCCTCCA	
	Frax-AF	GCTGCCAGGGGGCGTGCGGCA	
Msh2	Msh2wt1[v]	GCTCACTTAGATGCCATTGT	(REITMAIR *et al.* 1995)
	Msh2wt2[v]	AAAGTGCACGTCATTTGGA	
	Msh2mut[§]	GCCTTCTTGACGAGTTCTTC	
Msh6	MO10	TGGAAGGATTGGAGCTACGG	(EDELMANN *et al.* 1997)
	MO11	TTACCTCCTCCACTGACGTG	
	MO12	CAAGCCCCTTTCTTTGTTTG	
	MO13	ACCACTTCCTCATCCCTGG	
PolB	Y265-F1	AGAAAAGCAGCTTCCAGCAG	(SENEJANI *et al.* 2012)
	Y265-R4	CAGACTTTCCAAGTGCAGGAT	

*= This primer was labeled with 6-carboxyfluorescein (FAM). [v]= Wt specific, [§]= KO-

Table 6.2 PCR conditions used for genotyping.

PCR	Reaction Mix		Cycling conditions		Fragment size
CGG	Reagents	PCR reaction	Steps	Conditions	~480bp
	2 x KAPA master mix	12 µl	I	95°C; 3 min	
	DMSO	0.5 µl	II	95°C; 15 sec	
	Betaine 5M	10 µl	III	65°C; 15 sec	
	Frax-m4-Hex	0.25 µl	IV	72°C; 15 sec	
	Frax-m5	0.25 µl	V	72°C; 10 min	
			(step 2-4 x **35 cycles)**		
Msh2	Reagents	PCR reaction	Steps	Conditions	460bp (KO) 164bp (WT)
	H2O	18 µl	I	95°C; 1min	
	5x-MyTaq master mix	5 µl	II	95°C; 15 sec	
	Msh2-wt	0.25 µl	III	55°C; 30 sec	
	Msh2-wt2	0.25 µl	IV	72°C; 10 sec	
	Msh2-mut	0.25 µl	V	72°C; 10 min	
	MyTaq enzyme	0.25 µl	(step 2-4 x **5 cycles)**		
Msh6	Reagents	PCR reaction	Steps	Conditions	400bp (KO) 560bp (WT)
	H2O	8 µl	I	95°C; 3 min	
	2 x KAPA master mix	10 µl	II	95°C; 15 sec	
	Msh6-M010	0.25 µl	III	60°C; 15 sec	
	Msh6-M011	0.25 µl	IV	72°C; 15 sec	
	Msh6-M012	0.25 µl	V	72°C; 10 min	
	Msh6-M013	0.25 µl	(step 2-4 x **30cycles)**		
PolB	Reagents	PCR reaction	Steps	Conditions	500bp (Mut) 450bp (WT)
	H2O	18 µl	I	95°C; 1min	
	5x-MyTaq master mix	5 µl	II	95°C; 15 sec	
	Msh2-wt	0.25 µl	III	55°C; 30 sec	
	Msh2-wt2	0.25 µl	IV	72°C; 10 sec	
	Msh2-mut	0.25 µl	V	72°C; 10 min	
	MyTaq enzyme	0.25 µl	(step 2-4 x **30cycles)**		

bp = base pairs, I = initial denaturation, II = denaturation, III = annealing, IV= elongation, V = extension.

6.4 Repeat analysis

In order to determine the repeat number in each animal, the CGG•CCG repeat-tract was amplified using ~100 ng of genomic DNA from tail or mouse tissues by PCR as described above for the PM genotyping of the animals except that the Frax m4 primer was labeled with either 6-carboxyfluorescein (FAM) or 4, 7, 2′, 4′, 5′, 7′-hexachloro-6-carboxyfluorescein (HEX). Two microliters of the PCR product was mixed with 10 µl HiDi Formamide and 0.5 µl of GeneScan LIZ 600™, LIZ1200™ or Rox 500™ size standard (Applied Biosystems, Carlsbad, CA). The mixture was denatured at 95°C for 3 minutes and chilled immediately on ice, then resolved by capillary electrophoresis on an ABI Genetic Analyzer (Applied Biosystems). The PCR profiles were analyzed using GeneMapper Software v4.0

For analysis of intergenerational instability, the repeat size in the tail DNA taken at weaning (3 weeks of age) from each animal was compared to the repeat size of the tail DNA of the transmitting parent taken at the same age to determine whether expansion or contraction had occurred. In the case of somatic instability analysis, the repeat size in adult mouse tissue at the time of euthanasia was compared to the repeat size of the same animal in the tail taken at weaning. To quantify the degree of expansion in various organs of our mice, the somatic instability index from different organs was computed as previously described (LEE *et al.* 2010) with some modification. Briefly, the highest peak in the Gene Mapper profile of each sample analyzed was identified and peaks with heights less than 10% of this peak were excluded from further analysis. “The peak height was standardized by dividing the peak height of each peak

by the sum of the heights of all peaks. The normalized peak heights were then multiplied by the changes from the original allele, and these values were summed to generate the instability index"(LEE *et al.* 2010; HAMLIN *et al.* 2011). The change in repeat length of each peak was calculated from the highest peak in the tail sample determined in the mouse at 3 weeks of age (main allele), or in the heart sample determined at time of death since we have found that somatic instability is minimal in heart, and thus the main allele in the heart of an adult mouse is similar to the main allele in the tail taken at 3 weeks

6.5 The XCI assay

We developed a methylation–sensitive PCR assay to distinguish the repeat profile present on PM alleles located on active X chromosomes from those on inactive ones. The assay is based on the fact that the Frax m5 primer, one of the primers used for the genotyping of the PM allele and determination of repeat number in the FX PM PCR, contains a recognition site for *Sau96I*, a restriction enzyme sensitive to methylation. Therefore, digestion of DNA with *Sau96I* will eliminate PM alleles found on the active X chromosome.

10 U of *Sau96I* restriction enzyme was used to digest one microgram (1 µg) of genomic DNA in a total volume of 50 µl for 4 hours at 37°C in 1x buffer 4 (New England Biolabs, Ipswich, MA). For each sample an undigested control was similarly prepared without the enzyme. To monitor the extent of digestion, 500 ng of pUC19, a plasmid that has 6 *Sau96I* recognition sites was added to each genomic DNA sample. Samples were

monitored by agarose gel electrophoresis to confirm complete digestion of the plasmid. The reaction mixes were then heated at 65°C for 20 minutes to inactivate the enzyme. Following DNA digestion, 2 µl (~100 ng) of the digest and undigested DNA was used in the FX PM PCR assay. Two microliters of the PCR product was then mixed with 10 µl HiDi Formamide and 0.5 µl of LIZ1200™ size standard. The mixture was denatured at 95°C for 3 minutes and chilled immediately on ice, and the PCR products were analyzed as described above.

6.6 X inactivation ratio assay

Since amplification across the repeat tract is not quantitative, we needed a different assay to determine the degree of XCI skewing. Therefore, we designed a different methylation-sensitive PCR based assay that take advantage of the cleavage site for the restriction enzyme *EaeI* found in the neomycin marker that was part of the construct used to generate the FX PM mouse line (ENTEZAM *et al.* 2007). This marker is located in intron 1 and DNA digestion by *EaeI* will occur only if the template is located on the active X chromosome.

One microgram of genomic DNA was incubated in 1x CutSmart buffer (New England Biolabs, Ipswich, MA) with 10 U of the restriction enzyme *EaeI* (New England Biolabs) 4 hours at 37°C. The enzyme was heat inactivated for 20 minutes at 65°C. Following DNA digestion, quantitative real-time PCR (qRT-PCR) was performed using power SYBR master mix (Applied Biosystems) in a real time PCR machine (StepOne Plus, Applied Biosystems). qRT-PCR reactions were performed in 96-well plates

(MicroAmp fast # 4346907, Applied Biosystems, Carlsbad, CA) in triplicate. A final volume of 20 µl master mix was prepared containing 10 µl of 2x power SYBR master mix (Applied Biosystems, Carlsbad, CA), 0.4 µl of 10 µM XCI-TF and XCI-TR primers (Table 3), 7.2 µl of water and 2 µl of the digested DNA (or undigested DNA).

A control reaction for *EaeI* restriction enzyme digestion was also included. The reaction was done using primer pairs XCI-CF and XCI-CR that amplifies a region just upstream of the target amplicon (*EaeI* sensitive amplicon). This region has an *EaeI* site that does not contain CpG residues and is consequently digested whether the region is methylated or not. The *EaeI*-sensitive PCR (target) and control master mixes were prepared separately and added to the 96 wells reaction plate (MicroAmp Fast, Applied Biosystems) with an automatic repetitive pipette (Rainin AutoRep E, Rainin, Oakland, CA), followed by adding the digested or undigested DNA. The real time plate was then sealed with a plate sealer (MicroAmp, Applied Biosystems, Carlsbad, CA). The mixtures were then mixed gently and bubbles removed by gently tapping the bottom of the wells. Subsequently, the plate was briefly spun down at 1000 rpm to bring all the liquid to the bottom of the wells. The PCR conditions are provided in Table 4. The yield of the *EaeI* digested methylation sensitive amplicons relative to undigested material was calculated using the comparative Ct method (SCHMITTGEN AND LIVAK 2008).

Table 6.3 Primers used in XCI methylation assay.

PCR	Primer ID	Sequence (5'-3')	Fragment size
EaeI- target	XCI-TF	TGGACGAAGAGCATCAGGGG	268bp
	XCI-TR	GCGATACCGTAAAGCACGAG	
EaeI- Control	XCI-CF	CGGTTCTTTTTGTCAAGACCGA	195bp
	XCI-CR	CAGGAGCAAGGTGAGATGACA	

T= target, C= control, F= forward, R= reverse, bp= base pair

Table 6.4 qRT-PCR conditions for XCI assay

PCR name	**Reaction Mix**		**Cycling conditions**	
	Reagent	**Vol (µl)**	**Step**	**Conditions**
Target/control	H_2O	7.2 µl	I	95°C; 3 min
	2x power SYBR mix	10 µl	II	95°C; 15 sec
	XCI-F/XCI-CF	0.4 µl	III	60°C; 15 sec
	XCI-TR/ XCI-CR	0.4 µl	IV	72°C; 15 sec
			V	72°C; 10 min
			(step 2-4 x **30cycles)**	

I= initial denaturation, II= denaturation, III= annealing, IV= elongation
V= extension, vol= volume

6.7 Western blotting

6.7.1 Protein extraction and quantification

~ 20 mg of mouse tissues were emulsified using a tissue homogenizer (Precellys 24, Bertin Technologies, Rockville, MD) with T-PER (total lysate) or NE-PER (nuclear fraction) protein extraction reagents (Pierce Biotechnology, Rockford, IL) supplemented with phosphatase inhibitor cocktail-3 (Sigma-Aldrich, St. Louis, MO) and protease inhibitor cocktail (Roche Applied Science, Indianapolis, IN) according to the data sheet of the product and the manufacturer ‘s instructions. Bradford assay, a colorimetric protein assay (REISNER *et al.* 1975), was used to determine the protein concentration using either a Bio-Rad protein assay kit (cat # 500-0205; Bio-Rad, Hercules, CA) or a Coomassie Plus-The Better Bradford assay kit (cat # 1856268; Pierce Biotechnology, Rockford, IL) according to the manufacturer’s recommendations.

Briefly, using the Bio-Rad kit, one milliliter of Bradford reagent was mixed with 5 µl of the sample in a plastic disposable cuvette, incubated for 5 minutes at room temperature and absorbance read at 595 nm in a spectrophotometer. To measure the protein concentration using the Pierce kit, 1.5 ml of Bradford reagent was mixed with 5 µl of samples in a disposable cuvette as well and incubated for 10 minutes at room temperature. A calibration curve was generated each time a protein assay was performed with Bovine Serum Albumin (BSA) dilutions of known concentrations.

6.7.2 Polyacrylamide gel electrophoresis and Western blotting

Polyacrylamide gel electrophoresis of SDS-denatured proteins was performed using the Novex NuPAGE SDS-PAGE gel system, a high-performance denaturing polyacrylamide gel electrophoresis system (Life Technologies, Carlsbad, CA). Twenty to 30 µg proteins were resolved on either a 3-8% Novex Tris-Acetate protein or 4-12% Novex Bis-Tris gels (Life Technologies Carlsbad, CA) and transferred onto nitrocellulose membranes using 1X NuPAGE Transfer Buffer (Life Technologies) at 100 V, at room temperature, for 1 hour. 10 % of methanol was added to the transfer buffer. After the protein transfer onto the membranes, the membranes were blocked for 1 hour in 5% ECL Prime blocking agent (GE Healthcare Bio-Sciences, Pittsburgh, PA) in 1xTBST (10 mM Tris-HCl pH 7.5, 0.15 mM NaCl and 0.01% Tween 20) at room temperature, then probed with different antibodies as listed in table 5 overnight at 4°C. The membranes were then washed 3 times in 1xTBST (10 mM Tris-HCl pH 7.5, 0.15 mM NaCl and 0.01% Tween 20). After incubation with the appropriate secondary antibody (anti-rabbit or anti-mouse), the membranes were once again washed 3 times in 1xTBS. The FluorChem M imaging system (Proteinsimple, Santa Clara, CA) was used to detect the protein on the membrane. Prior to protein detection, ECL Prime detection reagent (GE Healthcare Bio-Sciences) was added onto the membranes.

Table 6.5 Antibodies used for Western blots

Antibody ID	Species/clonality/isotype	Dilution	Source	Catalog #
Anti-APE1	Rabbit polyclonal IgG	1:10,000	Abcam (Cambridge, MA, USA)	Ab137708
Anti-ATRX	Rabbit polyclonal IgG	1:1,000	Abcam (Cambridge, MA, USA)	Ab97508
Anti-DNAligase 1	Rabbit polyclonal IgG	1:5,000	Abcam (Cambridge, MA, USA)	Ab76232
Anti-DNA ligase 3	Mouse monoclonal IgG	1:3,000	Biosciences (Franklin Lakes, NJ, USA)	BD611876
Anti-FEN1	Rabbit polyclonal IgG	1:10,000	Abcam (Cambridge, MA, USA)	Ab70815
Anti-MSH2	Rabbit polyclonal IgG	1:10,000	Abcam (Cambridge, MA, USA)	Ab70270
Anti-MSH3	Mouse monoclonal IgG	1:1,000	Santa Cruz (Dallas, TX, USA)	Sc-271079
Anti-MSH6	Mouse monoclonal IgG	1:500	Biosciences (Franklin Lakes, NJ, USA)	BD610918
Anti-NEIL1	Rabbit polyclonal IgG	1:2,000	Abcam (Cambridge, MA, USA)	Ab21337
Anti-OGG1	Rabbit polyclonal IgG	1:3,000	Proteintech Group (Chicago, IL, USA)	15125-1-AP
Anti-Polδ	Rabbit polyclonal IgG	1:10,000	Abcam (Cambridge, MA, USA)	Ab186407
Anti-Polε	Rabbit polyclonal IgG	1:10,000	Abcam (Cambridge, MA, USA)	Ab74308
Anti-Polβ	Rabbit polyclonal IgG	1:1,000	Abcam (Cambridge, MA, USA)	Ab26343
Anti-SMC1	Mouse monoclonal IgG	1:1,000	Abcam (Cambridge, MA, USA)	Ab52324
Anti-Ube1a/b	Rabbit polyclonal IgG	1:1,000	Cell Signaling (Boston, MA, USA)	4891

6.8 ELISA essay:

ELISA essays were done using the MSH3 ELISA kit (cat # MBS9326679, MyBioSource, San Diego, CA) according to the manufacturer's recommendations. Briefly, mouse tissues were emulsified using a tissue homogenizer (Precellys 24, Bertin Technologies, Rockville, MD) in 1X PBS. The homogenates were then spun down for 15' at 500g. The supernatant is then collected. Fifty microliters of standards of known concentrations are added in standard wells and 50 µl of each sample added to sample wells of the 96 wells ELISA plate pre-coated with MSH3. One hundred microliters of HRP-conjugate reagent was added to each well and the mixture incubated for 60 minutes at 37˚C. The liquid was then drained off and the plate washed 4 times with the wash solution provided in the kit. Chromatogen solutions (solution A and B provided in the kit) were then added to all wells followed by 15 minutes incubation in the dark. The plate was then read at 450 nm using a Model 680 microplate reader (Bio-Rad Laboratories).

6.9 Statistical analysis

A web version of the GraphPad QuickCalcs Software (http://www.graphpad.com) was used to perform statistical evaluation of our data. . Fisher's exact test was used in the evaluation of the frequency of expansions, contractions and unchanged alleles. Comparison of continuous data sets was done by the Student's t test. In some circumstances, the 95% confidence interval (95% CI) was estimated for the amount of expansions, contractions and unchanged alleles using the GraphPad implementation of

the modified Wald method (http://www.graphpad.com). All calculated *P*-values are two-tailed and a *P*-value of <0.05 was considered as significant.

6.10 List of publications

Lokanga, R.A., Senejani, A.G., Sweasy, J.B. and Usdin, K. (2015) Heterozygosity for a hypomorphic *PolB* mutation reduces the expansion frequency in a mouse model of the Fragile X-related disorders. *Plos Genet*, **11**, e10051

Lokanga, R.A, Zhao, X.N., Entezam, A. and Usdin, K. (2014) X inactivationplays a major role in the gender bias in somatic expansion in a mouse model of the fragile X-related disorders: implications for the mechanism of repeat expansion. *Hum Mol Genet*, **23**, 4985-4994

Lokanga, R.A., Zhao, X.N. and Usdin, K. (2014) The mismatch repair protein MSH2 is rate limiting for repeat expansion in a fragile X premutation mouse model. *Human mutation*, **35**, 129-136

Lokanga, R.A., Entezam, A., Kumari, D., Yudkin, D., Qin, M., Smith, C.B. and Usdin, K. (2013) Somatic expansion in mouse and human carriers of fragile X premutation alleles. *Human mutation*, **34**, 1

Lokanga, R.A., Zhao, X.N., Kimaada A., Kumari, D., and Usdin, K. (2016) MutSα contributes significantly to repeat expansion in a Fragile X premutation mouse model but only in the presence of MutSβ (manuscript in revision)

6.11 References

Abdel-Rahman, W. M., J. P. Mecklin and P. Peltomaki, 2006 The genetics of HNPCC: application to diagnosis and screening. Crit Rev Oncol Hematol 58**:** 208-220.

Adams, J. S., P. E. Adams, D. Nguyen, J. A. Brunberg, F. Tassone *et al.*, 2007 Volumetric brain changes in females with fragile X-associated tremor/ataxia syndrome (FXTAS). Neurology 69**:** 851-859.

Allen, D., D. C. Herbert, C. A. McMahan, V. Rotrekl, R. W. Sobol *et al.*, 2008 Mutagenesis is elevated in male germ cells obtained from DNA polymerase-beta heterozygous mice. Biol Reprod 79**:** 824-831.

Allingham-Hawkins, D. J., R. Babul-Hirji, D. Chitayat, J. J. Holden, K. T. Yang *et al.*, 1999 Fragile X premutation is a significant risk factor for premature ovarian failure: the International Collaborative POF in Fragile X study--preliminary data. Am J Med Genet 83**:** 322-325.

Amouroux, R., A. Campalans, B. Epe and J. P. Radicella, 2010 Oxidative stress triggers the preferential assembly of base excision repair complexes on open chromatin regions. Nucleic Acids Res 38**:** 2878-2890.

Ashley, C. T., J. S. Sutcliffe, C. B. Kunst, H. A. Leiner, E. E. Eichler *et al.*, 1993 Human and murine FMR-1: alternative splicing and translational initiation downstream of the CGG-repeat. Nat Genet 4**:** 244-251.

Ashley-Koch, A. E., H. Robinson, A. E. Glicksman, S. L. Nolin, C. E. Schwartz *et al.*, 1998 Examination of factors associated with instability of the FMR1 CGG repeat. Am J Hum Genet 63**:** 776-785.

Aspinwall, R., D. G. Rothwell, T. Roldan-Arjona, C. Anselmino, C. J. Ward *et al.*, 1997 Cloning and characterization of a functional human homolog of Escherichia coli endonuclease III. Proc Natl Acad Sci U S A 94**:** 109-114.

Avner, P., and E. Heard, 2001 X-chromosome inactivation: counting, choice and initiation. Nat Rev Genet 2**:** 59-67.

Bakker, C. E., Y. de Diego Otero, C. Bontekoe, P. Raghoe, T. Luteijn *et al.*, 2000 Immunocytochemical and biochemical characterization of FMRP, FXR1P, and FXR2P in the mouse. Exp Cell Res 258**:** 162-170.

Bear, M., 2005 The mGluR theory of fragile X mental retardation. Neuropsychopharmacology 30**:** S59-S59.

Beard, W. A., and S. H. Wilson, 2006 Structure and mechanism of DNA polymerase Beta. Chem Rev 106**:** 361-382.

Bennetto, L., B. F. Pennington, D. Porter, A. K. Taylor and R. J. Hagerman, 2001 Profile of cognitive functioning in women with the fragile X mutation. Neuropsychology 15**:** 290-299.

Berry-Kravis, E., K. Potanos, D. Weinberg, L. Zhou and C. G. Goetz, 2005 Fragile X-associated tremor/ataxia syndrome in sisters related to X-inactivation. Ann Neurol 57**:** 144-147.

Bontekoe, C. J., C. E. Bakker, I. M. Nieuwenhuizen, H. van der Linde, H. Lans *et al.*, 2001 Instability of a (CGG)98 repeat in the Fmr1 promoter. Hum Mol Genet 10**:** 1693-1699.

Bourn, R. L., I. De Biase, R. M. Pinto, C. Sandi, S. Al-Mahdawi *et al.*, 2012 Pms2 suppresses large expansions of the (GAA.TTC)n sequence in neuronal tissues. PLoS One 7**:** e47085.

Brega, A. G., G. Goodrich, R. E. Bennett, D. Hessl, K. Engle *et al.*, 2008 The primary cognitive deficit among males with fragile X-associated tremor/ataxia syndrome (FXTAS) is a dysexecutive syndrome. J Clin Exp Neuropsychol 30**:** 853-869.

Brown, V., K. Small, L. Lakkis, Y. Feng, C. Gunter *et al.*, 1998 Purified recombinant Fmrp exhibits selective RNA binding as an intrinsic property of the fragile X mental retardation protein. J Biol Chem 273**:** 15521-15527.

Brunberg, J. A., S. Jacquemont, R. J. Hagerman, E. M. Berry-Kravis, J. Grigsby *et al.*, 2002 Fragile X premutation carriers: Characteristic MR imaging findings of adult male patients with progressive cerebellar and cognitive dysfunction. American Journal of Neuroradiology 23**:** 1757-1766.

Butler, M. G., R. Pratesi, M. S. Watson, W. R. Breg and D. N. Singh, 1993 Anthropometric and craniofacial patterns in mentally retarded males with emphasis on the fragile X syndrome. Clin Genet 44**:** 129-138.

Carethers, J. M., M. Koi and S. S. Tseng-Rogenski, 2015 EMAST is a Form of Microsatellite Instability That is Initiated by Inflammation and Modulates Colorectal Cancer Progression. Genes 6**:** 185-205.

Chan, N. L. S., J. Z. Guo, T. Y. Zhang, G. G. Mao, C. X. Hou *et al.*, 2013 Coordinated Processing of 3 ' Slipped (CAG)(n)/(CTG)(n) Hairpins by DNA Polymerases beta and delta Preferentially Induces Repeat Expansions. Journal of Biological Chemistry 288**:** 15015-15022.

Chen, L. S., F. Tassone, P. Sahota and P. J. Hagerman, 2003 The (CGG)n repeat element within the 5' untranslated region of the FMR1 message provides both positive and negative cis effects on in vivo translation of a downstream reporter. Hum Mol Genet 12**:** 3067-3074.

Chong, S. S., A. E. McCall, J. Cota, S. H. Subramony, H. T. Orr *et al.*, 1995 Gametic and somatic tissue-specific heterogeneity of the expanded SCA1 CAG repeat in spinocerebellar ataxia type 1. Nat Genet 10**:** 344-350.

Clark, A. B., F. Valle, K. Drotschmann, R. K. Gary and T. A. Kunkel, 2000 Functional interaction of proliferating cell nuclear antigen with MSH2-MSH6 and MSH2-MSH3 complexes. J Biol Chem 275**:** 36498-36501.

Cohen, S., K. Masyn, J. Adams, D. Hessl, S. Rivera *et al.*, 2006 Molecular and imaging correlates of the fragile X-associated tremor/ataxia syndrome. Neurology 67**:** 1426-1431.

Comery, T. A., J. B. Harris, P. J. Willems, B. A. Oostra, S. A. Irwin *et al.*, 1997 Abnormal dendritic spines in fragile X knockout mice: maturation and pruning deficits. Proc Natl Acad Sci U S A 94**:** 5401-5404.

Cornish, K. M., F. Munir and G. Cross, 2001 Differential impact of the FMR-1 full mutation on memory and attention functioning : a neuropsychological perspective. J Cogn Neurosci 13**:** 144-150.

De Biase, I., A. Rasmussen, D. Endres, S. Al-Mahdawi, A. Monticelli *et al.*, 2007 Progressive GAA expansions in dorsal root ganglia of Friedreich's ataxia patients. Ann Neurol 61**:** 55-60.

De Boulle, K., A. J. Verkerk, E. Reyniers, L. Vits, J. Hendrickx *et al.*, 1993 A point mutation in the FMR-1 gene associated with fragile X mental retardation. Nat Genet 3**:** 31-35.

de Graaff, E., P. Rouillard, P. J. Willems, A. P. Smits, F. Rousseau *et al.*, 1995 Hotspot for deletions in the CGG repeat region of FMR1 in fragile X patients. Hum Mol Genet 4**:** 45-49.

de Vries, B. B., A. M. Wiegers, A. P. Smits, S. Mohkamsing, H. J. Duivenvoorden *et al.*, 1996 Mental status of females with an FMR1 gene full mutation. Am J Hum Genet 58**:** 1025-1032.

Devys, D., Y. Lutz, N. Rouyer, J. P. Bellocq and J. L. Mandel, 1993 The FMR-1 protein is cytoplasmic, most abundant in neurons and appears normal in carriers of a fragile X premutation. Nat Genet 4**:** 335-340.

Dianov, G. L., K. M. Sleeth, Dianova, II and S. L. Allinson, 2003 Repair of abasic sites in DNA. Mutat Res 531**:** 157-163.

Dragileva, E., A. Hendricks, A. Teed, T. Gillis, E. T. Lopez *et al.*, 2009 Intergenerational and striatal CAG repeat instability in Huntington's disease knock-in mice involve different DNA repair genes. Neurobiol Dis 33**:** 37-47.

Du, J., E. Campau, E. Soragni, S. Ku, J. W. Puckett *et al.*, 2012 Role of mismatch repair enzymes in GAA.TTC triplet-repeat expansion in Friedreich ataxia induced pluripotent stem cells. J Biol Chem 287**:** 29861-29872.

Edelmann, W., K. Yang, A. Umar, J. Heyer, K. Lau *et al.*, 1997 Mutation in the mismatch repair gene Msh6 causes cancer susceptibility. Cell 91**:** 467-477.

Eichler, E. E., C. B. Kunst, K. A. Lugenbeel, O. A. Ryder, D. Davison *et al.*, 1995 Evolution of the cryptic FMR1 CGG repeat. Nat Genet 11**:** 301-308.

Eichler, E. E., S. Richards, R. A. Gibbs and D. L. Nelson, 1993 Fine structure of the human FMR1 gene. Hum Mol Genet 2**:** 1147-1153.

Ennis, S., D. Ward and A. Murray, 2006 Nonlinear association between CGG repeat number and age of menopause in FMR1 premutation carriers. European Journal of Human Genetics 14**:** 253-255.

Entezam, A., R. Biacsi, B. Orrison, T. Saha, G. E. Hoffman *et al.*, 2007 Regional FMRP deficits and large repeat expansions into the full mutation range in a new Fragile X premutation mouse model. Gene 395**:** 125-134.

Entezam, A., A. R. Lokanga, W. Le, G. Hoffman and K. Usdin, 2010 Potassium bromate, a potent DNA oxidizing agent, exacerbates germline repeat expansion in a fragile X premutation mouse model. Hum Mutat 31**:** 611-616.

Entezam, A., and K. Usdin, 2008 ATR protects the genome against CGG.CCG-repeat expansion in Fragile X premutation mice. Nucleic Acids Res 36**:** 1050-1056.

Entezam, A., and K. Usdin, 2009 ATM and ATR protect the genome against two different types of tandem repeat instability in Fragile X premutation mice. Nucleic Acids Res 37**:** 6371-6377.

Ezzatizadeh, V., R. M. Pinto, C. Sandi, M. Sandi, S. Al-Mahdawi *et al.*, 2012 The mismatch repair system protects against intergenerational GAA repeat instability in a Friedreich ataxia mouse model. Neurobiol Dis 46**:** 165-171.

Ezzatizadeh, V., C. Sandi, M. Sandi, S. Anjomani-Virmouni, S. Al-Mahdawi *et al.*, 2014 MutLalpha heterodimers modify the molecular phenotype of Friedreich ataxia. PLoS One 9**:** e100523.

Feng, Y., F. P. Zhang, L. K. Lokey, J. L. Chastain, L. Lakkis *et al.*, 1995 Translational Suppression by Trinucleotide Repeat Expansion at Fmr1. Science 268**:** 731-734.

Fisch, G. S., K. Snow, S. N. Thibodeau, M. Chalifaux, J. J. Holden *et al.*, 1995 The fragile X premutation in carriers and its effect on mutation size in offspring. Am J Hum Genet 56**:** 1147-1155.

Flores-Rozas, H., and R. D. Kolodner, 1998 The Saccharomyces cerevisiae MLH3 gene functions in MSH3-dependent suppression of frameshift mutations. Proc Natl Acad Sci U S A 95**:** 12404-12409.

Foiry, L., L. Dong, C. Savouret, L. Hubert, H. te Riele *et al.*, 2006 Msh3 is a limiting factor in the formation of intergenerational CTG expansions in DM1 transgenic mice. Hum Genet 119**:** 520-526.

Fortune, M. T., C. Vassilopoulos, M. I. Coolbaugh, M. J. Siciliano and D. G. Monckton, 2000 Dramatic, expansion-biased, age-dependent, tissue-specific somatic mosaicism in a transgenic mouse model of triplet repeat instability. Hum Mol Genet 9**:** 439-445.

Friedberg, E. C., and T. Lindahl, 2004 Inroads into base excision repair II. The discovery of DNA glycosylases. "An N-glycosidase from Escherichia coli that releases free uracil from DNA containing deaminated cytosine residues," Proc. Nat. Acad. Sci. USA, 1974. DNA Repair (Amst) 3**:** 1532-1536; discussion 1531-1532.

Fu, X., D. Zheng, J. Liao, Q. Li, Y. Lin *et al.*, 2015 Alternatively spliced products lacking exon 12 dominate the expression of fragile X mental retardation 1 gene in human tissues. Mol Med Rep 12**:** 1957-1962.

Fu, Y. H., D. P. A. Kuhl, A. Pizzuti, M. Pieretti, J. S. Sutcliffe *et al.*, 1991 Variation of the Cgg Repeat at the Fragile-X Site Results in Genetic Instability - Resolution of the Sherman Paradox. Cell 67**:** 1047-1058.

Gacy, A. M., G. Goellner, N. Juranic, S. Macura and C. T. McMurray, 1995 Trinucleotide repeats that expand in human disease form hairpin structures in vitro. Cell 81**:** 533-540.

Gacy, A. M., G. M. Goellner, C. Spiro, X. Chen, G. Gupta *et al.*, 1998 GAA instability in Friedreich's Ataxia shares a common, DNA-directed and intraallelic mechanism with other trinucleotide diseases. Mol Cell 1**:** 583-593.

Gatchel, J. R., and H. Y. Zoghbi, 2005 Diseases of unstable repeat expansion: Mechanisms and common principles. Nature Reviews Genetics 6**:** 743-755.

Gibson, T. J., P. M. Rice, J. D. Thompson and J. Heringa, 1993 KH domains within the FMR1 sequence suggest that fragile X syndrome stems from a defect in RNA metabolism. Trends Biochem Sci 18**:** 331-333.

Glaser, D., D. Wohrle, U. Salat, W. Vogel, P. Steinbach *et al.*, 1999 Mitotic behavior of expanded CGG repeats studied on cultured cells: Further evidence for methylation-mediated triplet repeat stability in fragile X syndrome. American Journal of Medical Genetics 84**:** 226-228.

Gomes-Pereira, M., M. T. Fortune, L. Ingram, J. P. McAbney and D. G. Monckton, 2004 Pms2 is a genetic enhancer of trinucleotide CAG.CTG repeat somatic mosaicism: implications for the mechanism of triplet repeat expansion. Hum Mol Genet 13**:** 1815-1825.

Gonitel, R., H. Moffitt, K. Sathasivam, B. Woodman, P. J. Detloff *et al.*, 2008 DNA instability in postmitotic neurons. Proc Natl Acad Sci U S A 105**:** 3467-3472.

Goula, A. V., B. R. Berquist, D. M. Wilson, V. C. Wheeler, Y. Trottier *et al.*, 2009 Stoichiometry of Base Excision Repair Proteins Correlates with Increased Somatic CAG Instability in Striatum over Cerebellum in Huntington's Disease Transgenic Mice. Plos Genetics 5.

Goula, A. V., A. Stys, J. P. Chan, Y. Trottier, R. Festenstein *et al.*, 2012 Transcription elongation and tissue-specific somatic CAG instability. PLoS Genet 8**:** e1003051.

Grabczyk, E., and K. Usdin, 1999 Generation of microgram quantities of trinucleotide repeat tracts of defined length, interspersion pattern, and orientation. Anal Biochem 267**:** 241-243.

Grasso, M., E. M. Boon, S. Filipovic-Sadic, P. A. van Bunderen, E. Gennaro *et al.*, 2014 A novel methylation PCR that offers standardized determination of FMR1 methylation and CGG repeat length without southern blot analysis. J Mol Diagn 16**:** 23-31.

Greco, C. M., R. J. Hagerman, F. Tassone, A. E. Chudley, M. R. Del Bigio *et al.*, 2002 Neuronal intranuclear inclusions in a new cerebellar tremor/ataxia syndrome among fragile X carriers. Brain 125**:** 1760-1771.

Greco, C. M., K. Soontrapornchai, J. Wirojanan, J. E. Gould, P. J. Hagerman *et al.*, 2007 Testicular and pituitary inclusion formation in fragile X associated tremor/ataxia syndrome. J Urol 177**:** 1434-1437.

Hagerman, P., 2013 Fragile X-associated tremor/ataxia syndrome (FXTAS): pathology and mechanisms. Acta Neuropathol 126**:** 1-19.

Hagerman, R., and P. Hagerman, 2013 Advances in clinical and molecular understanding of the FMR1 premutation and fragile X-associated tremor/ataxia syndrome. Lancet Neurol 12**:** 786-798.

Hagerman, R. J., B. R. Leavitt, F. Farzin, S. Jacquemont, C. M. Greco *et al.*, 2004 Fragile-X-associated tremor/ataxia syndrome (FXTAS) in females with the FMR1 premutation. American Journal of Human Genetics 74**:** 1051-1056.

Hagerman, R. J., M. Leehey, W. Heinrichs, F. Tassone, R. Wilson *et al.*, 2001 Intention tremor, parkinsonism, and generalized brain atrophy in male carriers of fragile X. Neurology 57**:** 127-130.

Hamlin, A., Y. Liu, D. V. Nguyen, F. Tassone, L. Zhang *et al.*, 2011 Sleep apnea in fragile X premutation carriers with and without FXTAS. Am J Med Genet B Neuropsychiatr Genet 156B**:** 923-928.

Hamlin, A. A., D. Sukharev, L. Campos, Y. Mu, F. Tassone *et al.*, 2012 Hypertension in FMR1 premutation males with and without fragile X-associated tremor/ataxia syndrome (FXTAS). Am J Med Genet A 158A**:** 1304-1309.

Hammond, L. S., M. M. Macias, J. C. Tarleton and G. Shashidhar Pai, 1997 Fragile X syndrome and deletions in FMR1: new case and review of the literature. Am J Med Genet 72**:** 430-434.

Hatton, D. D., J. Sideris, M. Skinner, J. Mankowski, D. B. Bailey, Jr. *et al.*, 2006 Autistic behavior in children with fragile X syndrome: prevalence, stability, and the impact of FMRP. Am J Med Genet A 140A**:** 1804-1813.

Hecht, F., and G. R. Sutherland, 1985 Detection of fragile sites on human chromosomes. Clin Genet 28**:** 95-96.

Heitz, D., D. Devys, G. Imbert, C. Kretz and J. L. Mandel, 1992 Inheritance of the fragile X syndrome: size of the fragile X premutation is a major determinant of the transition to full mutation. J Med Genet 29**:** 794-801.

Huang, T., L. Y. Li, Y. Shen, X. B. Qin, Z. L. Pang *et al.*, 1996 Alternative splicing of the FMR1 gene in human fetal brain neurons. Am J Med Genet 64**:** 252-255.

Ishiguro, H., K. Yamada, H. Sawada, K. Nishii, N. Ichino *et al.*, 2001 Age-dependent and tissue-specific CAG repeat instability occurs in mouse knock-in for a mutant Huntington's disease gene. J Neurosci Res 65**:** 289-297.

Iwahashi, C. K., D. H. Yasui, H. J. An, C. M. Greco, F. Tassone *et al.*, 2006 Protein composition of the intranuclear inclusions of FXTAS. Brain 129**:** 256-271.

Iyer, R. R., A. Pluciennik, V. Burdett and P. L. Modrich, 2006 DNA mismatch repair: functions and mechanisms. Chem Rev 106**:** 302-323.

Jacquemont, S., R. J. Hagerman, M. Leehey, J. Grigsby, L. Zhang *et al.*, 2003 Fragile X premutation tremor/ataxia syndrome: Molecular, clinical, and neuroimaging correlates. American Journal of Human Genetics 72**:** 869-878.

Jin, P., R. Duan, A. Qurashi, Y. Qin, D. Tian *et al.*, 2007 Pur alpha binds to rCGG repeats and modulates repeat-mediated neurodegeneration in a Drosophila model of fragile X tremor/ataxia syndrome. Neuron 55**:** 556-564.

Kirchgessner, C. U., S. T. Warren and H. F. Willard, 1995 X inactivation of the FMR1 fragile X mental retardation gene. J Med Genet 32**:** 925-929.

Klungland, A., and T. Lindahl, 1997 Second pathway for completion of human DNA base excision-repair: reconstitution with purified proteins and requirement for DNase IV (FEN1). EMBO J 16**:** 3341-3348.

Kolodner, R., 1996 Biochemistry and genetics of eukaryotic mismatch repair. Genes Dev 10**:** 1433-1442.

Kolodner, R. D., and G. T. Marsischky, 1999 Eukaryotic DNA mismatch repair. Curr Opin Genet Dev 9**:** 89-96.

Kovtun, I. V., Y. Liu, M. Bjoras, A. Klungland, S. H. Wilson *et al.*, 2007 OGG1 initiates age-dependent CAG trinucleotide expansion in somatic cells. Nature 447**:** 447-452.

Kovtun, I. V., and C. T. McMurray, 2001 Trinucleotide expansion in haploid germ cells by gap repair. Nat Genet 27**:** 407-411.

Ku, S., E. Soragni, E. Campau, E. A. Thomas, G. Altun *et al.*, 2010 Friedreich's ataxia induced pluripotent stem cells model intergenerational GAATTC triplet repeat instability. Cell Stem Cell 7**:** 631-637.

Kudo, M., K. Sugasawa, T. Hori, T. Enomoto, F. Hanaoka *et al.*, 1991 Human ubiquitin-activating enzyme (E1): compensation for heat-labile mouse E1 and its gene localization on the X chromosome. Exp Cell Res 192**:** 110-117.

Kumari, D., R. Lokanga, D. Yudkin, X. N. Zhao and K. Usdin, 2012 Chromatin changes in the development and pathology of the Fragile X-associated disorders and Friedreich ataxia. Biochimica Et Biophysica Acta-Gene Regulatory Mechanisms 1819**:** 802-810.

Kunkel, T. A., and D. A. Erie, 2005 DNA mismatch repair. Annu Rev Biochem 74**:** 681-710.

La Spada, A. R., E. M. Wilson, D. B. Lubahn, A. E. Harding and K. H. Fischbeck, 1991 Androgen receptor gene mutations in X-linked spinal and bulbar muscular atrophy. Nature 352**:** 77-79.

Lalioti, M. D., H. S. Scott, C. Buresi, C. Rossier, A. Bottani *et al.*, 1997 Dodecamer repeat expansion in cystatin B gene in progressive myoclonus epilepsy. Nature 386**:** 847-851.

Lang, W. H., J. E. Coats, J. Majka, G. L. Hura, Y. Lin *et al.*, 2011 Conformational trapping of mismatch recognition complex MSH2/MSH3 on repair-resistant DNA loops. Proc Natl Acad Sci U S A 108**:** E837-844.

Latham, G. J., J. Coppinger, A. G. Hadd and S. L. Nolin, 2014 The role of AGG interruptions in fragile X repeat expansions: a twenty-year perspective. Front Genet 5**:** 244.

Lavedan, C., E. Grabczyk, K. Usdin and R. L. Nussbaum, 1998 Long uninterrupted CGG repeats within the first exon of the human FMR1 gene are not intrinsically unstable in transgenic mice. Genomics 50**:** 229-240.

Lavedan, C. N., L. Garrett and R. L. Nussbaum, 1997 Trinucleotide repeats (CGG)22TGG(CGG)43TGG(CGG)21 from the fragile X gene remain stable in transgenic mice. Hum Genet 100**:** 407-414.

Lee, J. M., J. Zhang, A. I. Su, J. R. Walker, T. Wiltshire *et al.*, 2010 A novel approach to investigate tissue-specific trinucleotide repeat instability. BMC Syst Biol 4**:** 29.

Li, G. M., 2008 Mechanisms and functions of DNA mismatch repair. Cell Res 18**:** 85-98.

Lia, A. S., H. Seznec, H. Hofmann-Radvanyi, F. Radvanyi, C. Duros *et al.*, 1998 Somatic instability of the CTG repeat in mice transgenic for the myotonic dystrophy region is age dependent but not correlated to the relative intertissue transcription levels and proliferative capacities. Hum Mol Genet 7**:** 1285-1291.

Lindahl, T., 1974 An N-glycosidase from Escherichia coli that releases free uracil from DNA containing deaminated cytosine residues. Proc Natl Acad Sci U S A 71**:** 3649-3653.

Liquori, C. L., K. Ricker, M. L. Moseley, J. F. Jacobsen, W. Kress *et al.*, 2001 Myotonic dystrophy type 2 caused by a CCTG expansion in intron 1 of ZNF9. Science 293**:** 864-867.

Liu, Y., W. A. Beard, D. D. Shock, R. Prasad, E. W. Hou *et al.*, 2005 DNA polymerase beta and flap endonuclease 1 enzymatic specificities sustain DNA synthesis for long patch base excision repair. J Biol Chem 280**:** 3665-3674.

Liu, Y., R. Prasad, W. A. Beard, E. W. Hou, J. K. Horton *et al.*, 2009 Coordination between polymerase beta and FEN1 can modulate CAG repeat expansion. J Biol Chem 284**:** 28352-28366.

Liu, Y., and S. H. Wilson, 2012 DNA base excision repair: a mechanism of trinucleotide repeat expansion. Trends Biochem Sci 37**:** 162-172.

Loeb, L. A., and R. J. Monnat, Jr., 2008 DNA polymerases and human disease. Nat Rev Genet 9**:** 594-604.

Lokanga, R. A., A. Entezam, D. Kumari, D. Yudkin, M. Qin *et al.*, 2013 Somatic expansion in mouse and human carriers of fragile X premutation alleles. Hum Mutat 34**:** 157-166.

Lokanga, R. A., X. N. Zhao and K. Usdin, 2014 The mismatch repair protein MSH2 is rate limiting for repeat expansion in a fragile X premutation mouse model. Hum Mutat 35**:** 129-136.

Lu, A. L., X. Li, Y. Gu, P. M. Wright and D. Y. Chang, 2001 Repair of oxidative DNA damage: mechanisms and functions. Cell Biochem Biophys 35**:** 141-170.

Lynch, H. T., W. Kimberling, W. A. Albano, J. F. Lynch, K. Biscone *et al.*, 1985 Hereditary nonpolyposis colorectal cancer (Lynch syndromes I and II). I. Clinical description of resource. Cancer 56**:** 934-938.

Lyon, M. F., 1961 Gene action in the X-chromosome of the mouse (Mus musculus L.). Nature 190**:** 372-373.

Lyons, J. I., G. R. Kerr and P. W. Mueller, 2015 Fragile X Syndrome: Scientific Background and Screening Technologies. J Mol Diagn 17**:** 463-471.

Malter, H. E., J. C. Iber, R. Willemsen, E. de Graaff, J. C. Tarleton *et al.*, 1997 Characterization of the full fragile X syndrome mutation in fetal gametes. Nat Genet 15**:** 165-169.

Manley, K., J. Pugh and A. Messer, 1999a Instability of the CAG repeat in immortalized fibroblast cell cultures from Huntington's disease transgenic mice. Brain Res 835**:** 74-79.

Manley, K., T. L. Shirley, L. Flaherty and A. Messer, 1999b Msh2 deficiency prevents in vivo somatic instability of the CAG repeat in Huntington disease transgenic mice. Nat Genet 23**:** 471-473.

Martin, J. P., and J. Bell, 1943 A Pedigree of Mental Defect Showing Sex-Linkage. J Neurol Psychiatry 6**:** 154-157.

Mason, A. G., S. Tome, J. P. Simard, R. T. Libby, T. K. Bammler *et al.*, 2014 Expression levels of DNA replication and repair genes predict regional somatic repeat instability in the brain but are not altered by polyglutamine disease protein expression or age. Hum Mol Genet 23**:** 1606-1618.

Matsuura, T., T. Yamagata, D. L. Burgess, A. Rasmussen, R. P. Grewal *et al.*, 2000 Large expansion of the ATTCT pentanucleotide repeat in spinocerebellar ataxia type 10. Nat Genet 26**:** 191-194.

McCulloch, S. D., L. Gu and G. M. Li, 2003 Bi-directional processing of DNA loops by mismatch repair-dependent and -independent pathways in human cells. J Biol Chem 278**:** 3891-3896.

McMurray, C. T., 2008 Hijacking of the mismatch repair system to cause CAG expansion and cell death in neurodegenerative disease. DNA Repair (Amst) 7**:** 1121-1134.

McMurray, C. T., 2010 Mechanisms of trinucleotide repeat instability during human development. Nat Rev Genet 11**:** 786-799.

Mirkin, S. M., 2006 DNA structures, repeat expansions and human hereditary disorders. Curr Opin Struct Biol 16**:** 351-358.

Mirkin, S. M., 2007 Expandable DNA repeats and human disease. Nature 447**:** 932-940.

Mitas, M., A. Yu, J. Dill and I. S. Haworth, 1995 Hairpin Properties of Single-Stranded-DNA Containing G+C-Rich Triplet Repeats - (Ctg)(15) and (Cgg)(15). Faseb Journal 9**:** A1324-A1324.

Modrich, P., and R. Lahue, 1996 Mismatch repair in replication fidelity, genetic recombination, and cancer biology. Annu Rev Biochem 65**:** 101-133.

Mollersen, L., A. D. Rowe, J. L. Illuzzi, G. A. Hildrestrand, K. J. Gerhold *et al.*, 2012 Neil1 is a genetic modifier of somatic and germline CAG trinucleotide repeat instability in R6/1 mice. Hum Mol Genet 21**:** 4939-4947.

Morales, F., J. M. Couto, C. F. Higham, G. Hogg, P. Cuenca *et al.*, 2012 Somatic instability of the expanded CTG triplet repeat in myotonic dystrophy type 1 is a heritable quantitative trait and modifier of disease severity. Hum Mol Genet 21**:** 3558-3567.

Morales, F., M. Vasquez, P. Cuenca, D. Campos, C. Santamaria *et al.*, 2015 Parental age effects, but no evidence for an intrauterine effect in the transmission of myotonic dystrophy type 1. Eur J Hum Genet 23**:** 646-653.

Nolin, S. L., W. T. Brown, A. Glicksman, G. E. Houck, Jr., A. D. Gargano *et al.*, 2003 Expansion of the fragile X CGG repeat in females with premutation or intermediate alleles. Am J Hum Genet 72**:** 454-464.

Nolin, S. L., A. Glicksman, X. Ding, N. Ersalesi, W. T. Brown *et al.*, 2011 Fragile X analysis of 1112 prenatal samples from 1991 to 2010. Prenat Diagn 31**:** 925-931.

Nolin, S. L., S. Sah, A. Glicksman, S. L. Sherman, E. Allen *et al.*, 2013 Fragile X AGG analysis provides new risk predictions for 45-69 repeat alleles. American Journal of Medical Genetics Part A 161A**:** 771-778.

Oberle, I., F. Rousseau, D. Heitz, C. Kretz, D. Devys *et al.*, 1991 Instability of a 550-base pair DNA segment and abnormal methylation in fragile X syndrome. Science 252**:** 1097-1102.

Oh, S. Y., F. He, A. Krans, M. Frazer, J. P. Taylor *et al.*, 2015 RAN translation at CGG repeats induces ubiquitin proteasome system impairment in models of fragile X-associated tremor ataxia syndrome. Hum Mol Genet 24**:** 4317-4326

Ohzeki, S., A. Tachibana, K. Tatsumi and T. Kato, 1997 Spectra of spontaneous mutations at the hprt locus in colorectal carcinoma cell lines defective in mismatch repair. Carcinogenesis 18**:** 1127-1133.

Oostra, B. A., and R. Willemsen, 2009 FMR1: a gene with three faces. Biochim Biophys Acta 1790**:** 467-477.

Owen, B. A., Z. Yang, M. Lai, M. Gajec, J. D. Badger, 2nd *et al.*, 2005 (CAG)(n)-hairpin DNA binds to Msh2-Msh3 and changes properties of mismatch recognition. Nat Struct Mol Biol 12**:** 663-670.

Pearson, C. E., K. Nichol Edamura and J. D. Cleary, 2005 Repeat instability: mechanisms of dynamic mutations. Nat Rev Genet 6**:** 729-742.

Pearson, C. E., and R. R. Sinden, 1996 Alternative structures in duplex DNA formed within the trinucleotide repeats of the myotonic dystrophy and fragile X loci. Biochemistry 35**:** 5041-5053.

Pena-Diaz, J., S. Bregenhorn, M. Ghodgaonkar, C. Follonier, M. Artola-Boran *et al.*, 2012 Noncanonical mismatch repair as a source of genomic instability in human cells. Mol Cell 47**:** 669-680.

Pieretti, M., F. P. Zhang, Y. H. Fu, S. T. Warren, B. A. Oostra *et al.*, 1991 Absence of expression of the FMR-1 gene in fragile X syndrome. Cell 66**:** 817-822.

Pimentel, M. M., 1999 Fragile X syndrome (review). Int J Mol Med 3**:** 639-645.

Pinto, R. M., E. Dragileva, A. Kirby, A. Lloret, E. Lopez *et al.*, 2013 Mismatch repair genes Mlh1 and Mlh3 modify CAG instability in Huntington's disease mice: genome-wide and candidate approaches. PLoS Genet 9**:** e1003930.

Poynter, J. N., K. D. Siegmund, D. J. Weisenberger, T. I. Long, S. N. Thibodeau *et al.*, 2008 Molecular characterization of MSI-H colorectal cancer by MLHI promoter methylation, immunohistochemistry, and mismatch repair germline mutation screening. Cancer Epidemiol Biomarkers Prev 17**:** 3208-3215.

Pretto, D. I., G. Mendoza-Morales, J. Lo, R. Cao, A. Hadd *et al.*, 2014 CGG allele size somatic mosaicism and methylation in FMR1 premutation alleles. J Med Genet.

Primerano, B., F. Tassone, R. J. Hagerman, P. Hagerman, F. Amaldi *et al.*, 2002 Reduced FMR1 mRNA translation efficiency in fragile X patients with premutations. RNA 8**:** 1482-1488.

Rahner, N., V. Steinke, B. Schlegelberger, F. Eisinger, P. Hutter *et al.*, 2013 Clinical utility gene card for: Lynch syndrome (MLH1, MSH2, MSH6, PMS2, EPCAM) - update 2012. Eur J Hum Genet 21.

Reisner, A. H., P. Nemes and C. Bucholtz, 1975 Use of Coomassie Brilliant Blue G250 Perchloric-Acid Solution for Staining in Electrophoresis and Isoelectric Focusing on Polyacrylamide Gels. Analytical Biochemistry 64**:** 509-516.

Reitmair, A. H., R. Schmits, A. Ewel, B. Bapat, M. Redston *et al.*, 1995 MSH2 deficient mice are viable and susceptible to lymphoid tumours. Nat Genet 11**:** 64-70.

Renoux, A. J., and P. K. Todd, 2012 Neurodegeneration the RNA way. Prog Neurobiol 97**:** 173-189.

Richardson, T. E., A. E. Yu, Y. Wen, S. H. Yang and J. W. Simpkins, 2012 Estrogen prevents oxidative damage to the mitochondria in Friedreich's ataxia skin fibroblasts. PLoS One 7**:** e34600.

Rivera, C. M., B. R. Grossardt, D. J. Rhodes and W. A. Rocca, 2009 Increased mortality for neurological and mental diseases following early bilateral oophorectomy. Neuroepidemiology 33**:** 32-40.

Robertson, A. B., A. Klungland, T. Rognes and I. Leiros, 2009 DNA repair in mammalian cells: Base excision repair: the long and short of it. Cell Mol Life Sci 66**:** 981-993.

Rocca, W. A., B. R. Grossardt, V. M. Miller, L. T. Shuster and R. D. Brown, Jr., 2012 Premature menopause or early menopause and risk of ischemic stroke. Menopause 19**:** 272-277.

Rousseau, F., D. Heitz, V. Biancalana, S. Blumenfeld, C. Kretz *et al.*, 1991 Direct diagnosis by DNA analysis of the fragile X syndrome of mental retardation. N Engl J Med 325**:** 1673-1681.

Sancar, A., L. A. Lindsey-Boltz, K. Unsal-Kacmaz and S. Linn, 2004 Molecular mechanisms of mammalian DNA repair and the DNA damage checkpoints. Annu Rev Biochem 73**:** 39-85.

Santucci-Darmanin, S., and V. Paquis-Flucklinger, 2003 [Homologs of MutS and MutL during mammalian meiosis]. Med Sci (Paris) 19**:** 85-91.

Savouret, C., E. Brisson, J. Essers, R. Kanaar, A. Pastink *et al.*, 2003 CTG repeat instability and size variation timing in DNA repair-deficient mice. Embo Journal 22**:** 2264-2273.

Schmittgen, T. D., and K. J. Livak, 2008 Analyzing real-time PCR data by the comparative C(T) method. Nat Protoc 3**:** 1101-1108.

Scott, E. L., Q. G. Zhang, R. K. Vadlamudi and D. W. Brann, 2014 Premature menopause and risk of neurological disease: basic mechanisms and clinical implications. Mol Cell Endocrinol 389**:** 2-6.

Sellier, C., F. Freyermuth, R. Tabet, T. Tran, F. He *et al.*, 2013 Sequestration of DROSHA and DGCR8 by expanded CGG RNA repeats alters microRNA processing in fragile X-associated tremor/ataxia syndrome. Cell Rep 3**:** 869-880.

Sellier, C., F. Rau, Y. Liu, F. Tassone, R. K. Hukema *et al.*, 2010 Sam68 sequestration and partial loss of function are associated with splicing alterations in FXTAS patients. EMBO J 29**:** 1248-1261.

Senejani, A. G., S. Dalal, Y. Liu, T. P. Nottoli, J. M. McGrath *et al.*, 2012 Y265C DNA polymerase beta knockin mice survive past birth and accumulate base excision repair intermediate substrates. Proc Natl Acad Sci U S A 109**:** 6632-6637.

Seriola, A., C. Spits, J. P. Simard, P. Hilven, P. Haentjens *et al.*, 2011 Huntington's and myotonic dystrophy hESCs: down-regulated trinucleotide repeat instability and mismatch repair machinery expression upon differentiation. Human Molecular Genetics 20**:** 176-185.

Sherman, S. L., 2000 Premature ovarian failure in the fragile X syndrome. Am J Med Genet 97**:** 189-194.

Sherman, S. L., N. E. Morton, P. A. Jacobs and G. Turner, 1984 The marker (X) syndrome: a cytogenetic and genetic analysis. Ann Hum Genet 48**:** 21-37.

Siomi, H., M. C. Siomi, R. L. Nussbaum and G. Dreyfuss, 1993 The protein product of the fragile X gene, FMR1, has characteristics of an RNA-binding protein. Cell 74**:** 291-298.

Snow, K., D. J. Tester, K. E. Kruckeberg, D. J. Schaid and S. N. Thibodeau, 1994 Sequence analysis of the fragile X trinucleotide repeat: implications for the origin of the fragile X mutation. Hum Mol Genet 3**:** 1543-1551.

Spiro, C., R. Pelletier, M. L. Rolfsmeier, M. J. Dixon, R. S. Lahue *et al.*, 1999 Inhibition of FEN-1 processing by DNA secondary structure at trinucleotide repeats. Mol Cell 4**:** 1079-1085.

Sullivan, S. D., C. Welt and S. Sherman, 2011 FMR1 and the continuum of primary ovarian insufficiency. Semin Reprod Med 29**:** 299-307.

Swami, M., A. E. Hendricks, T. Gillis, T. Massood, J. Mysore *et al.*, 2009 Somatic expansion of the Huntington's disease CAG repeat in the brain is associated with an earlier age of disease onset. Hum Mol Genet 18**:** 3039-3047.

Tanaka, F., M. F. Reeves, Y. Ito, M. Matsumoto, M. Li *et al.*, 1999 Tissue-specific somatic mosaicism in spinal and bulbar muscular atrophy is dependent on CAG-repeat length and androgen receptor--gene expression level. Am J Hum Genet 65**:** 966-973.

Tassone, F., J. Adams, E. M. Berry-Kravis, S. S. Cohen, A. Brusco *et al.*, 2007a CGG repeat length correlates with age of onset of motor signs of the fragile X-associated tremor/ataxia syndrome (FXTAS). American Journal of Medical Genetics Part B-Neuropsychiatric Genetics 144B**:** 566-569.

Tassone, F., A. Beilina, C. Carosi, S. Albertosi, C. Bagni *et al.*, 2007b Elevated FMR1 mRNA in premutation carriers is due to increased transcription. Rna-a Publication of the Rna Society 13**:** 555-562.

Tassone, F., C. Iwahashi and P. J. Hagerman, 2004 FMR1 RNA within the intranuclear inclusions of fragile X-associated tremor/ataxia syndrome (FXTAS). RNA Biol 1**:** 103-105.

Telenius, H., B. Kremer, Y. P. Goldberg, J. Theilmann, S. E. Andrew *et al.*, 1994 Somatic and gonadal mosaicism of the Huntington disease gene CAG repeat in brain and sperm. Nat Genet 6**:** 409-414.

Terracciano, A., M. G. Pomponi, G. M. Marino, P. Chiurazzi, M. M. Rinaldi *et al.*, 2004 Expansion to full mutation of a FMR1 intermediate allele over two generations. Eur J Hum Genet 12**:** 333-336.

Thornton, C. A., K. Johnson and R. T. Moxley, 3rd, 1994 Myotonic dystrophy patients have larger CTG expansions in skeletal muscle than in leukocytes. Ann Neurol 35**:** 104-107.

Tian, L., C. Hou, K. Tian, N. C. Holcomb, L. Gu *et al.*, 2009 Mismatch recognition protein MutSbeta does not hijack (CAG)n hairpin repair in vitro. J Biol Chem 284**:** 20452-20456.

Tome, S., I. Holt, W. Edelmann, G. E. Morris, A. Munnich *et al.*, 2009 MSH2 ATPase domain mutation affects CTG*CAG repeat instability in transgenic mice. PLoS Genet 5**:** e1000482.

Tome, S., K. Manley, J. P. Simard, G. W. Clark, M. M. Slean *et al.*, 2013 MSH3 polymorphisms and protein levels affect CAG repeat instability in Huntington's disease mice. PLoS Genet 9**:** e1003280.

Tseng-Rogenski, S. S., H. Chung, M. B. Wilk, S. Zhang, M. Iwaizumi *et al.*, 2012 Oxidative Stress Induces Nuclear-to-Cytosol Shift of hMSH3, a Potential Mechanism for EMAST in Colorectal Cancer Cells. Plos One 7.

Ueno, S., K. Kondoh, Y. Kotani, O. Komure, S. Kuno *et al.*, 1995 Somatic mosaicism of CAG repeat in dentatorubral-pallidoluysian atrophy (DRPLA). Hum Mol Genet 4**:** 663-666.

Usdin, K., 1998 NGG-triplet repeats form similar intrastrand structures: implications for the triplet expansion diseases. Nucleic Acids Research 26**:** 4078-4085.

Usdin, K., 2008 The biological effects of simple tandem repeats: lessons from the repeat expansion diseases. Genome Res 18**:** 1011-1019.

Usdin, K., B. E. Hayward, D. Kumari, R. A. Lokanga, N. Sciascia *et al.*, 2014 Repeat-mediated genetic and epigenetic changes at the FMR1 locus in the Fragile X-related disorders. Front Genet 5**:** 226.

Usdin, K., N. C. House and C. H. Freudenreich, 2015 Repeat instability during DNA repair: Insights from model systems. Crit Rev Biochem Mol Biol 50**:** 142-167.

Usdin, K., and K. J. Woodford, 1995 CGG repeats associated with DNA instability and chromosome fragility form structures that block DNA synthesis in vitro. Nucleic Acids Res 23**:** 4202-4209.

van den Broek, W. J., M. R. Nelen, D. G. Wansink, M. M. Coerwinkel, H. te Riele *et al.*, 2002 Somatic expansion behaviour of the (CTG)n repeat in myotonic dystrophy knock-in mice is differentially affected by Msh3 and Msh6 mismatch-repair proteins. Hum Mol Genet 11**:** 191-198.

van den Broek, W. J., D. G. Wansink and B. Wieringa, 2007 Somatic CTG*CAG repeat instability in a mouse model for myotonic dystrophy type 1 is associated with changes in cell nuclearity and DNA ploidy. BMC Mol Biol 8**:** 61.

Verheij, C., C. E. Bakker, E. de Graaff, J. Keulemans, R. Willemsen *et al.*, 1993 Characterization and localization of the FMR-1 gene product associated with fragile X syndrome. Nature 363**:** 722-724.

Verkerk, A. J., E. de Graaff, K. De Boulle, E. E. Eichler, D. S. Konecki *et al.*, 1993 Alternative splicing in the fragile X gene FMR1. Hum Mol Genet 2**:** 399-404.

Verkerk, A. J. M. H., M. Pieretti, J. S. Sutcliffe, Y. H. Fu, D. P. A. Kuhl *et al.*, 1991 Identification of a Gene (Fmr-1) Containing a Cgg Repeat Coincident with a Breakpoint Cluster Region Exhibiting Length Variation in Fragile-X Syndrome. Cell 65**:** 905-914.

Walsh, P. S., N. J. Fildes and R. Reynolds, 1996 Sequence analysis and characterization of stutter products at the tetranucleotide repeat locus vWA. Nucleic Acids Res 24**:** 2807-2812.

Wheeler, V. C., L. A. Lebel, V. Vrbanac, A. Teed, H. te Riele *et al.*, 2003 Mismatch repair gene Msh2 modifies the timing of early disease in Hdh(Q111) striatum. Hum Mol Genet 12**:** 273-281.

Willemsen, R., C. Bontekoe, F. Tamanini, H. Galjaard, A. Hoogeveen *et al.*, 1996 Association of FMRP with ribosomal precursor particles in the nucleolus. Biochem Biophys Res Commun 225**:** 27-33.

Yang, F., T. Babak, J. Shendure and C. M. Disteche, 2010 Global survey of escape from X inactivation by RNA-sequencing in mouse. Genome Res 20**:** 614-622.

Yang, G. Z., S. J. Scherer, S. S. Shell, K. Yang, M. Kim *et al.*, 2004 Dominant effects of an Msh6 missense mutation on DNA repair and cancer susceptibility. Cancer Cell 6**:** 139-150.

Yrigollen, C. M., B. Durbin-Johnson, L. Gane, D. L. Nelson, R. Hagerman *et al.*, 2012 AGG interruptions within the maternal FMR1 gene reduce the risk of offspring with fragile X syndrome. Genet Med 14**:** 729-736.

Yrigollen, C. M., L. Martorell, B. Durbin-Johnson, M. Naudo, J. Genoves *et al.*, 2014 AGG interruptions and maternal age affect FMR1 CGG repeat allele stability during transmission. J Neurodev Disord 6**:** 24.

Yu, S., J. Mulley, D. Loesch, G. Turner, A. Donnelly *et al.*, 1992 Fragile-X syndrome: unique genetics of the heritable unstable element. Am J Hum Genet 50**:** 968-980.

Zhao, X. N., D. Kumari, S. Gupta, D. Wu, M. Evanitsky *et al.*, 2015 Mutsbeta generates both expansions and contractions in a mouse model of the Fragile X-associated disorders. Hum Mol Genet.

Zhao, X. N., and K. Usdin, 2014 Gender and cell-type-specific effects of the transcription-coupled repair protein, ERCC6/CSB, on repeat expansion in a mouse model of the fragile X-related disorders. Hum Mutat 35**:** 341-349.

Zhao, X. N., and K. Usdin, 2015 The Repeat Expansion Diseases: The dark side of DNA repair. DNA Repair (Amst) 32**:** 96-105.

Zhong, N., W. Ju, J. Pietrofesa, D. Wang, C. Dobkin *et al.*, 1996 Fragile X "gray zone" alleles: AGG patterns, expansion risks, and associated haplotypes. Am J Med Genet 64**:** 261-265.